ABSTRACT

Title of Dissertation: LOCALITY, SYMMETRY, AND DIGITAL SIMULATION OF QUANTUM SYSTEMS

Cong Minh Tran
Doctor of Philosophy, 2021

Dissertation Directed by: Professor Alexey V. Gorshkov
Department of Physics

Professor Jacob M. Taylor
Department of Physics

Besides potentially delivering a huge leap in computational power, quantum computers also offer an essential platform for simulating properties of quantum systems. Consequently, various algorithms have been developed for approximating the dynamics of a target system on quantum computers. But *generic* quantum simulation algorithms—developed to simulate all Hamiltonians—are unlikely to result in optimal simulations of most physically relevant systems; optimal quantum algorithms need to exploit unique properties of target systems. The aim of this dissertation is to study two prominent properties of physical systems, namely locality and symmetry, and subsequently leverage these properties to design efficient quantum simulation algorithms.

In the first part of the dissertation, we explore the locality of quantum systems and the fundamental limits on the propagation of information in power-law interacting systems. In particular, we prove upper limits on the speed at which information can propagate in power-law systems. We also demonstrate how such speed limits can be achieved

by protocols for transferring quantum information and generating quantum entanglement. We then use these speed limits to constrain the propagation of error and improve the performance of digital quantum simulation. Additionally, we consider the implications of the speed limits on entanglement generation, the dynamics of correlation, the heating time, and the scrambling time in power-law interacting systems.

In the second part of the dissertation, we propose a scheme to leverage the symmetry of target systems to suppress error in digital quantum simulation. We first study a phenomenon called *destructive error interference*, where the errors from different steps of a simulation cancel out each other. We then show that one can induce the destructive error interference by interweaving the simulation with unitary transformations generated by the symmetry of the target system, effectively providing a faster quantum simulation by symmetry protection. We also derive rigorous bounds on the error of a symmetry-protected simulation algorithm and identify conditions for optimal symmetry protection.

LOCALITY, SYMMETRY, AND DIGITAL SIMULATION OF QUANTUM SYSTEMS

by

Cong Minh Tran

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor Zohreh Davoudi, Chair
Professor Alexey V. Gorshkov, Co-Advisor
Professor Jacob M. Taylor, Co-Advisor
Professor Alicia J. Kollár
Professor Andrew M. Childs

Acknowledgments

This dissertation would not have been possible without my advisors, Alexey Gorshkov and Jacob Taylor. Their endless enthusiasm and patience have made this five-year journey look like a platter of Chesapeake blue crabs I had—I was not sure if I would finish it at first but it left me hungry for more in the end.

I have also been blessed with so many other great mentors and friends during this journey. I am grateful for the support from Andrew Childs, Guang Hao Low, and Yuan Su, from whom I have learned almost everything I know about quantum algorithms. I also appreciate the opportunities to work with Andrew Lucas, who taught me to be bold in research, Zhe-Xuan Gong, with whom I have had many inspiring discussions, and James Garrison, who have influenced how I write since the early stage of this Ph.D.

My undergraduate advisor, Tomasz Paterek, also holds a special place in my heart. Tomek was there when I first set foot in academia. He ignited my interest in quantum information science, taught me Bell inequalities, and walked me through writing my very first paper. I miss our casual lunch conversations and dearly wish we would meet again one day.

I would also like to thank my colleagues and friends with whom I have shared many great memories and discussions over the last five years: Christopher Baldwin, Aniruddha Bapat, Przemyslaw Bienias, Lucas Brady, Danniel Carney, Su-Kuan Chu, Abhinav Deshpande, Adam Ehrenberg, Zachary Eldredge, Michael Foss-Feig, Luis Pedro Garcia-Pintos, James Garrison, Andrew Guo, Yaroslav Kharkov, Simon Lieu, Nishad Maskara,

Marek Miller, Brittany Richman, Eddie Schoute, Yuan Su, Paraj Titum, Shuchen Zhu, and many others. I will particularly miss board-game nights and occasional birthday parties with Abhinav, Andrew, Ani, and Eddie.

Finally, I would like to thank my family for always being there for me. Nhung has walked with me and loved me every step of the way, even when there were an entire world between us. I also owe so much to my parents for their love and sacrifices for me. I miss them and I am sure they miss me everyday too.

Citations to Previously Published Work

Much of this dissertation has previously appeared in published papers. Here, we provide citations to those papers, in addition to related papers which are not presented in this disseration.

Chapter 2:

- "The Lieb-Robinson light cone for power-law interactions," M. C. Tran, A. Y. Guo, C. L. Baldwin, A. Ehrenberg, A. V. Gorshkov, A. Lucas, arXiv:2103.15828 (2021)
  - Further work related to the Lieb-Robinson bound in power-law interacting systems is presented in "Signaling and Scrambling with Strongly Long-Range Interactions," A. Y. Guo, M. C. Tran, A. M. Childs, A. V. Gorshkov, Z.-X. Gong, Phys. Rev. A 102, 010401 (2020).

Chapter 3:

- "Optimal State Transfer and Entanglement Generation in Power-law Interacting Systems," M. C. Tran, A. Y. Guo, A. Deshpande, A. Lucas, A. V. Gorshkov, arXiv: 2010.02930 (2020), to appear in Phys. Rev. X.

Chapter 4:

- "Hierarchy of linear light cones with long-range interactions," M. C. Tran, C.-F. Chen, A. Ehrenberg, A. Y. Guo, A. Deshpande, Y. Hong, Z.-X. Gong, A. V. Gorshkov, A. Lucas, Phys. Rev. X 10, 031009 (2020).

Chapter 5:

- "Locality and digital quantum simulation of power-law interactions," M. C. Tran, A. Y. Guo, Y. Su, J. R. Garrison, Z. Eldredge, M. Foss-Feig, A. M. Childs, A. V. Gorshkov, Phys. Rev. X 9, 031006 (2019).

  – Further work related to the simulation of local systems is presented in "Theory of Trotter Error with Commutator Scaling," A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, S. Zhu, Phys. Rev. X 11, 011020 (2021).

Chapter 6:

- "Hierarchy of linear light cones with long-range interactions," M. C. Tran, C.-F. Chen, A. Ehrenberg, A. Y. Guo, A. Deshpande, Y. Hong, Z.-X. Gong, A. V. Gorshkov, A. Lucas, Phys. Rev. X 10, 031009 (2020).

- "Locality and Heating in Periodically Driven, Power-law Interacting Systems," M. C. Tran, A. Ehrenberg, A. Y. Guo, P. Titum, D. A. Abanin, A. V. Gorshkov, Phys. Rev. A 100, 052103 (2019)

- 'Lieb-Robinson bounds on n-partite connected correlation functions," M. C. Tran, J. R. Garrison, Z.-X. Gong, A. V. Gorshkov, Phys. Rev. A 96, 052334 (2017)

  – Further work related to the simulation of bosonic systems is presented in "Dynamical phase transitions in sampling complexity," A. Deshpande, B. Fefferman, M. C. Tran, M. Foss-Feig, A. V. Gorshkov, Phys. Rev. Lett. 121, 030501 (2018).

– Further work related to the simulation of bosonic systems with power-law interactions is presented in "Complexity phase diagram for interacting and long-range bosonic Hamiltonians," N. Maskara, A. Deshpande, A. Ehrenberg, M. C. Tran, B. Fefferman, A. V. Gorshkov, arXiv:1906.04178 (2019).

Chapter 7:

- "Destructive Error Interference in Product-Formula Lattice Simulation," M. C. Tran, S.-K. Chu, Y. Su, A. M. Childs, A. V. Gorshkov, Phys. Rev. Lett. 124, 220502 (2020).

Chapter 8:

- "Faster Digital Quantum Simulation by Symmetry Protection," M. C. Tran, Y. Su, D. Carney, J. M. Taylor, PRX Quantum 2, 010323 (2021).

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Overview of the dissertation

Digital simulation is an important instrument in the discovery of physics, providing us with key insights into the dynamics of quantum systems. Instead of hours of clever mathematical manipulations, one could simply simulate the dynamics of the electrons in two molecules of different species as they are brought together in a chemical reaction and thereby learn the rate of reaction. One could also simulate the ground-state behaviors of electrons in a crystal and infer macroscopic properties of the corresponding material.

Due to quantum phenomena such as interference and entanglement, the simulations of quantum systems are often beyond the capability of classical computers. Quantum computers offer a solution in such situations. On a quantum computer, we would simulate the dynamics of the electrons by first mapping the electrons themselves onto "qubits"—the units of information on a quantum computer—and then mapping their dynamics onto a series of elementary operations ("quantum gates") applied to the qubits. Under these quantum gates, the qubits would mimic the dynamics of the electrons, allowing us to infer the properties of the electrons by measuring the qubits at the end of the simulation.

Finding the optimal mapping that minimizes the number of qubits, the number of elementary operations, and the error of the simulation is the central goal of quantum

simulation algorithms. In Sec. 1.2 below, we will provide a brief review of an important class of quantum simulation algorithms: the product formulas. The product formulas do not assume any special structure of the Hamiltonian and can be applied to simulate most quantum systems. This feature, however, is also a caveat; the product formulas may not provide any advantages, compared to the simulations of generic systems, even when we supply them with advanced knowledge about the target systems. The aim of this dissertation is to answer a fundamental question: Can we design better quantum simulation algorithms by exploiting some common properties of quantum systems?

In Part I of this dissertation, we will focus on using the locality of quantum systems to speed up the digital quantum simulation. Most physical interactions decay with distance as a power law $1/r^\alpha$, where r is the distance and $\alpha > 0$ is a constant. Examples include the Coulomb interaction ($\alpha = 1$), the dipole-dipole interaction ($\alpha = 3$), the Van der Waals interaction ($\alpha = 6$), and the nearest-neighbor interaction ($\alpha \to \infty$). The fact that these interactions decay with distance implies a notion of locality in the systems: it takes time for information to propagate in quantum systems. In Chapter 2, we will discuss the Lieb-Robinson bound, which provides an upper bound on the speed at which information can propagate in power-law systems. We will then present in Chapter 3 a protocol that transfers quantum states and generates entanglement at the speed that saturates this bound. In Chapter 4, we will discuss a hierarchy of speed limits, besides the Lieb-Robinson bound, that constrain different information processing tasks. In Chapter 5, we will discuss how a constraint on the propagation of information also constrains the propagation of simulation error, enabling the design of a faster quantum simulation algorithm. Finally, in Chapter 6, we will present several other implications of the locality

on the growth of correlations, the transfer of quantum states, the complexity of boson sampling, the generation of topologically ordered states, the clustering of correlations, and the heating rate in power-law interacting systems.

Symmetry provides another valuable, yet under-explored, resource for speeding up digital quantum simulation. In Part II of this dissertation, we will discuss a technique that leverages the symmetry in the target systems to improve quantum simulation algorithms. In Chapter 7, we will study a phenomenon called *destructive error interference*, where the error from different steps of a simulation may cancel out each other, resulting in a small total error of the simulation. In Chapter 8, we will then discuss how we can use the symmetry of the target systems to induce the destructive interference, which suppresses the error of the simulation and effectively enables a faster simulation.

1.2 Digital Quantum Simulation

In quantum mechanics, given a system described by a Hamiltonian H, its dynamics are given by the unitary operator $U = e^{-iHt}$, where t is the time of the evolution. In general, this unitary U acts non-trivially on every qubit representing the system. Meanwhile, the elementary operations of a quantum computer are often *two-local*—each elementary operation affects at most two qubits at a time. Quantum simulation algorithms provide mappings between the evolution unitary U and such elementary operations.

The first quantum simulation algorithm proposed by Lloyd [1] uses the Trotter product formula, also known as the first-order product formula [2, 3]. Given a Hamiltonian $H = \sum_{\mu=1}^{L} h_\mu$ that is a sum of L interaction terms h_μ, where each acts non-trivially on at

most two qubits, the Trotter product formula decomposes the evolution under H into the evolutions under each interaction h_μ:

$$e^{-iHt} \approx e^{-ih_L t} \dots e^{-ih_2 t} e^{-ih_1 t} =: S_1(t). \tag{1.1}$$

This decomposition is exact if the terms h_μ mutually commute: $[h_\mu, h_\nu] \coloneqq h_\mu h_\nu - h_\nu h_\mu = 0$. Otherwise, the error of the decomposition is upper bounded by a sum over the norms of the commutators between the terms of the Hamiltonian:

$$\|e^{-iHt} - e^{-ih_L t} \dots e^{-ih_2 t} e^{-ih_1 t}\| \leq \sum_{\mu,\nu=1}^{L} \|[h_\mu, h_\nu]\| t^2, \tag{1.2}$$

where $\|\cdot\|$ refers to the operator norm—the maximum singular value of a matrix. The error is second order in t and is small for short-time simulations. To control the error in a long-time simulation, we typically divide the time t into a number of small time steps, which we call the Trotter number r, and apply the Trotter product formula to each time step:

$$e^{-iHt} = \left(e^{-iHt/r}\right)^r \approx \left(e^{-ih_L t/r} \dots e^{-ih_2 t/r} e^{-ih_1 t/r}\right)^r. \tag{1.3}$$

This strategy results in a total error not exceeding $\sum_{\mu,\nu=1}^{L} \|[h_\mu, h_\nu]\| t^2/r$, which decreases as we increase the Trotter number. Because increasing r also increases the number of quantum gates used in the simulation, to minimize the gate count, we would choose the minimal value of r such that the total error of the simulation meets a certain tolerance ε.

The Trotter product formula can be generalized to algorithms that approximate the

time evolution e^{-iHt} up to higher-order in time [4, 5, 6, 7]. For example, the pth-order Suzuki-Trotter product formulas, for even $p \geq 2$, are symmetrized versions of the Trotter formula and can be constructed recursively on p [2]:

$$S_2(t) \coloneqq e^{-ih_L t/2} \dots e^{-ih_2 t/2} e^{-ih_1 t/2} e^{-ih_1 t/2} e^{-ih_2 t/2} \dots e^{-ih_L t/2}, \tag{1.4}$$

$$S_p(t) \coloneqq S_{p-2}(\lambda_p t)^2 . S_{p-2}((1-4\lambda_p)t) . S_{p-2}(\lambda_p t)^2, \qquad (p > 2) \tag{1.5}$$

where $\lambda_p = 1/(4 - 4^{1/(p-1)})$. The pth-order Suzuki-Trotter product formulas approximate the time evolution up to the pth order in time:

$$\|e^{-iHt} - S_p(t)\| \leq \frac{2}{p+1} \Upsilon^{p+1} \alpha_{\text{comm}} t^{p+1}, \tag{1.6}$$

where $\Upsilon = 2 \times 5^{p/2-1}$ and α_{comm} is a sum over the norms of all p-layered nested commutators between the terms of the Hamiltonian:

$$\alpha_{\text{comm}} \coloneqq \sum_{\mu_0,\dots,\mu_p=1}^{L} \| \left[h_{\mu_p}, \dots [h_{\mu_1}, h_{\mu_0}] \dots \right] \|. \tag{1.7}$$

Besides the product formulas, more advanced quantum simulation algorithms have been developed, including those that are based on quantum walks, linear combinations of unitaries [8, 8, 9], and quantum signal processing [10, 11]. These algorithms asymptotically reduce the cost of digital quantum simulation in terms of the number of gates used in the limit of large time or large system size. Despite these developments, the product formulas, due to their simple construction, remain the most popular algorithms for near-

term implementations of digital quantum simulation. In particular, the small prefactor in the scaling of the gate count of the first-order product formula compared to the more advanced quantum simulation algorithms makes it more attractive for simulations where the evolution time and the system size are not too large [7].

Part I

Locality and Digital Simulation of Quantum Systems

Chapter 2: Locality in Quantum Systems

Over a century ago, Einstein realized that there is a speed limit to information propagation. If no physical object or signal can travel faster than light, then the speed of light itself must constrain the dynamics of quantum information and entanglement. In ordinary quantum systems, however, *emergent* speed limits can arise that place more stringent restrictions on information propagation than does the speed of light. For example, in quantum spin systems with nearest-neighbor interactions on a lattice, Lieb and Robinson proved in 1972 that there is a *finite velocity* of information propagation [12].

Of course, most non-relativistic physical systems realized in experiments include long-range interactions such as the Coulomb interaction, the dipole-dipole interaction, or the van-der-Waals interaction. Each of these decays with distance as a power law $1/r^\alpha$ for some exponent α. What is the fundamental speed limit on the propagation of quantum information in these systems?

Despite the importance of this question in designing and constraining the operation of future quantum technologies [13, 14, 15, 16, 17], bounding information propagation in systems with power-law interactions has been a notoriously challenging mathematical physics problem. In 2005, Hastings and Koma [18] showed that it takes a time $t \gtrsim \log r$ to send information a distance r, for all $\alpha > d$, where d is the dimension of the lattice.

By analogy to Einstein's relativity, we say that there is at least a "logarithmic light cone" for such power-law interactions. However, it was suspected that this bound was far from tight, and ten years later it was shown that $t \gtrsim r^{\gamma}$, for an exponent $0 < \gamma < 1$ when $\alpha > 2d$ [19, 20, 21]. In 2019, Chen and Lucas [22] proved the existence of a linear light cone ($t \gtrsim r$) for all $\alpha > 3$ in $d = 1$; Kuwahara and Saito [23] later generalized this result to higher dimensions, finding a linear light cone for all $\alpha > 2d + 1$. These recent results prove that power-law interactions are, for all practical purposes, entirely local for sufficiently large α.

A natural question is then how small α must be in order to *break* a linear light cone. Fast state-transfer and entanglement-generation protocols developed in the past year [23, 24, 25, 26] have ultimately demonstrated that the time t required to send information a distance r obeys $t \lesssim r^{\min(\alpha-2d,1)}$ for any $\alpha > 2d$ and $t \lesssim r^{o(1)}$ for $\alpha < d$, where $o(1)$ is an arbitrarily small constant. Combining all best known results in the literature leads to the diagram shown in Fig. 2.1, which compares known information-transfer protocols to corresponding Lieb-Robinson bounds.

In this Chapter, we complete this extensive literature on Lieb-Robinson bounds for power-law interactions [18, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33], by proving that quantum information is contained within the Lieb-Robinson light cone $t \gtrsim r^{\min(\alpha-2d-\varepsilon,1)}$, for any $\varepsilon > 0$. This result closes the remaining gap between bounds and protocols in Fig. 2.1, and concludes the fifteen-year quest to understand the fundamental speed limit on quantum information in the presence of power-law interactions. We sketch the proof of the result in the main text and refer readers to Appendix A for a rigorous treatment.

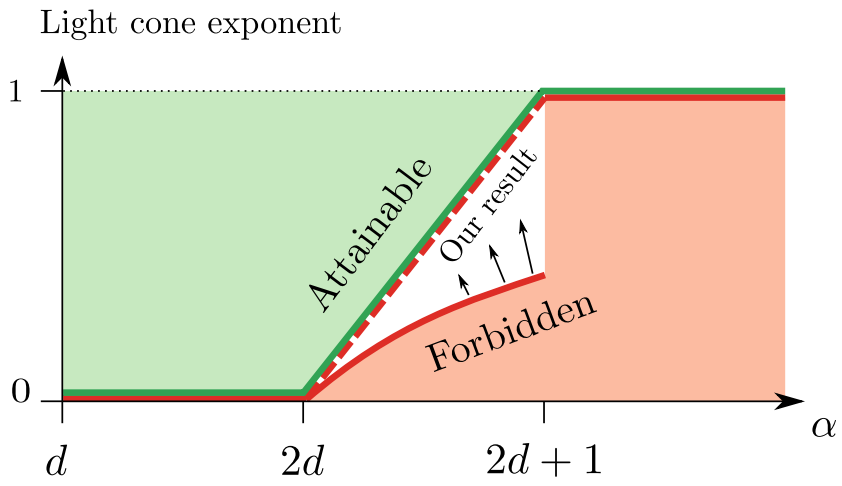

Figure 2.1: The gap in the Lieb-Robinson literature in $d > 1$ dimensions. The red solid lines represent the exponent γ of the Lieb-Robinson light cone $t \gtrsim r^\gamma$ in literature. The green solid lines correspond to the light cone exponents of best-known information-propagating protocols. Accordingly, the green region corresponds to attainable light cone exponents, whereas the red region is forbidden by the known bounds. Our result (red dashed line) closes the gap in our understanding of the Lieb-Robinson light cone.

2.1 The Lieb-Robinson Light Cone for Power-Law Interactions

We consider a d-dimensional regular lattice Λ, a finite-level system at every site of the lattice, and a two-body power-law Hamiltonian $H(t)$ with an exponent α supported on the lattice. Specifically, we assume $H(t) = \sum_{i,j\in\Lambda} h_{ij}(t)$ is a sum of two-body terms h_{ij} supported on sites i, j such that $\|h_{ij}(t)\| \leq 1/\mathrm{dist}(i,j)^\alpha$ for all $i \neq j$, where $\|\cdot\|$ is the operator norm and $\mathrm{dist}(i,j)$ is the distance between i, j. In the following discussion, we assume Λ is a hypercubic lattice of qubits for simplicity.

We use $\mathcal{L}$ to denote the Liouvillian corresponding to the Heisenberg evolution under Hamiltonian H, i.e. $\mathcal{L}\,|O) \equiv |i[H,O])$ for any operator O, and use $e^{\mathcal{L}t}\,|O) \equiv |O(t))$ to denote the time-evolved version of the operator O. We also use $\mathbb{P}_r^{(i)}\,|O)$ to denote an operator constructed from O by decomposing O into a sum of Pauli strings and removing

strings that are supported entirely within a ball of radius r from i. Colloquially speaking, $\mathbb{P}_r^{(i)}|O)$ is the component of O that has non-trivial support on sites a distance at least r from site i. If i is the origin of the lattice, we drop the superscript i and simply write $\mathbb{P}_r$ for brevity.

Given a unit-norm operator O initially supported at the origin, our main result is a bound on how much O spreads to a distance r and beyond under the evolution $e^{\mathcal{L}t}$:

Theorem 1. *For any $\alpha \in (2d, 2d+1)$ and an arbitrarily small $\varepsilon > 0$, there exist constants $c, C \geq 0$ such that*

$$\|\mathbb{P}_r e^{\mathcal{L}t}|O)\| \leq C\left(\frac{t}{r^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\frac{\varepsilon}{2}} \tag{2.1}$$

holds for all $1 \leq t \leq cr^{\alpha-2d-\varepsilon}$.

Because $\|\mathbb{P}_r e^{\mathcal{L}t}|O)\|$ can be both upper- and lower-bounded by linear functions of $\sup_A \|[A, e^{\mathcal{L}t}O]\|$, where A is a unit-norm operator supported at least a distance r from O, Eq. (2.1) is equivalent to a bound on the unequal-time commutators commonly used in the Lieb-Robinson literature.

For $\alpha \in (2d, 2d+1)$, by setting the left-hand side of Eq. (2.1) to a constant, Theorem 1 implies the light cone $t \gtrsim r^{\alpha-2d-\varepsilon}$ for some ε that can be made arbitrarily small. Note that our definition does not require $\|h_{ij}\|$ to decay exactly as $1/\text{dist}(i,j)^\alpha$; it may actually decay faster than $1/\text{dist}(i,j)^\alpha$ and still satisfy the condition of a power-law interaction with an exponent α. Therefore, for $\alpha \geq 2d+1$ and power-law Hamiltonians $H = \sum_{ij} h_{ij}$ satisfying $\|h_{ij}\| \leq 1/\text{dist}(i,j)^\alpha < 1/\text{dist}(i,j)^{2d+1-\varepsilon}$, Theorem 1 implies a linear light cone $t \gtrsim r^{1-2\varepsilon}$.

2.2 Sketch of Proof

For simplicity, we assume here that the lattice diameter is $\mathcal{O}(r)$. We show in Appendix A that interactions whose ranges are much larger than r do not contribute significantly to the dynamics of O and, therefore, can be safely removed from the Hamiltonian. Our strategy is to group the interactions of the Hamiltonian by their ranges, prove a bound for short-range interactions, and recursively add longer-range interactions to the Hamiltonian.

Specifically, we choose $\ell_k \equiv L^k$ for $k = 1, \ldots, n$, where L, n are parameters to be chosen later. We use H_k to denote those terms of H with range at most ℓ_k and use $\mathcal{L}_k \equiv i[H_k, \cdot]$ to denote the corresponding Liouvillian. We start with the standard Lieb-Robinson bound for H_1 [12]:

$$\|\mathbb{P}_r e^{\mathcal{L}_1 t}(O)\| \lesssim e^{\frac{v_1 t - r}{\ell_1}}, \tag{2.2}$$

where $v_1 \propto \ell_1 = L$ is the rescaled Lieb-Robinson velocity, and prove a bound for H_2 by adding $V_2 \equiv H_2 - H_1$, i.e., interactions of range between ℓ_1 and ℓ_2, to the Hamiltonian H_1.

For that, we move into the interaction picture of H_1 so that we can decompose the evolution $e^{\mathcal{L}_2 t} = e^{\mathcal{L}_{2,I} t} e^{\mathcal{L}_1 t}$ into two consecutive evolutions, where $e^{\mathcal{L}_{2,I} t}$ is the evolution under $V_{2,I} \equiv e^{\mathcal{L}_1 t} V_2$. Loosely speaking, the light cone induced by H_2 will be a "sum" of the light cones induced by H_1 and $V_{2,I}$ individually (see Appendix A for a proof.) With the light cone of H_1 given by Eq. (2.2), our task is to find the light cone of $V_{2,I}$.

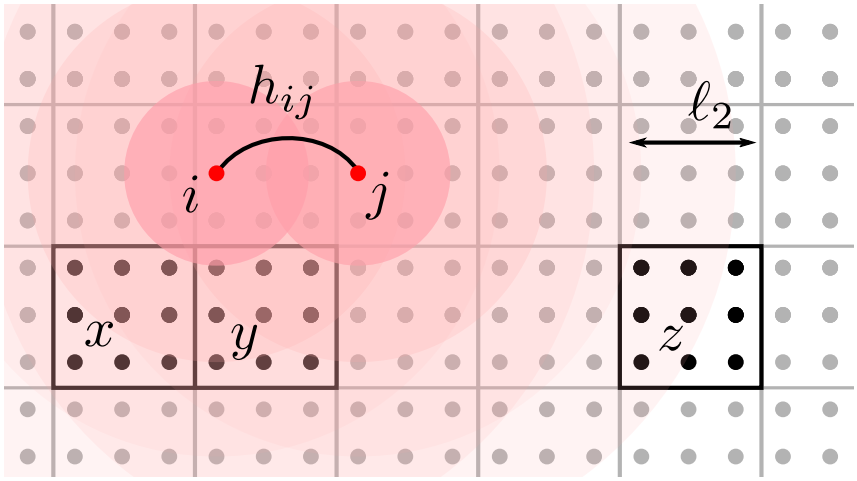

Figure 2.2: We study the structure of $V_{2,I}$ by dividing the lattice into hypercubes of length ℓ_2 (labeled by x, y, and z for example). In the interaction picture, how much each $e^{\mathcal{L}_1 t} h_{ij}$ contributes to the pair-wise "effective interaction" between two hypercubes depends on how strongly the support of $e^{\mathcal{L}_1 t} h_{ij}$ (represented by the shaded area) overlaps with the hypercubes. Because of the bound in Eq. (2.2), the evolved operator $e^{\mathcal{L}_1 t} h_{ij}$ is largely confined to the light cones induced by $\mathcal{L}_1$ around i and j (the smallest disks around i and j). The component of $e^{\mathcal{L}_1 t} h_{ij}$ supported outside this light cone is exponentially suppressed with distance (represented by lighter shade). Consequently, the effective interaction between the hypercubes x and z is exponentially smaller than the one between x and y.

For this purpose, we consider the structure of $V_{2,I}$ and show that, with a suitable rescaling of the lattice, the interactions in $V_{2,I}$ decay exponentially with distance. We then obtain the light cone of $V_{2,I}$ using the standard Lieb-Robinson bound on the rescaled lattice. Specifically, we divide the lattice into non-overlapping hypercubes of length ℓ_2 (see Fig. 2.2). Given x, y as the centers of two hypercubes, we define $\mathrm{dist}(x, y)/\ell_2$ to be the rescaled distance between the hypercubes. We shall estimate the strength of the interaction between hypercubes under the Hamiltonian $V_{2,I}$.

We first consider the case $t = 0$ so that $V_{2,I} = V_2$. Because each interaction in V_2 has range at most ℓ_2, no interaction h_{ij} is supported on two distinct hypercubes unless they are nearest neighbors. Therefore, only nearest-neighboring hypercubes may interact under $V_{2,I} = V_2$.

The case $t > 0$ is slightly more complicated. The support of an interaction h_{ij} in V_2 may expand under $e^{\mathcal{L}_1 t}$, and, hence, non-nearest-neighboring hypercubes may interact with each other. However, due to Eq. (2.2), the support of $e^{\mathcal{L}_1 t} h_{ij}$ would largely remain inside the balls of radius $v_1 t$ around i, j. The interactions between hypercubes are exponentially suppressed with distance by Eq. (2.2). Therefore, the system of hypercubes would interact via a nearly finite-range interaction (see Fig. 2.2).

To apply the standard Lieb-Robinson bound for this system of hypercubes, we estimate the maximum effective interaction between any pair of nearest-neighboring hypercubes centered on x, y. In particular, assuming $v_1 t \leq \ell_2$, the primary contributions to such an interaction come from $\propto \ell_2^d \times \ell_2^d = \ell_2^{2d}$ interaction terms $e^{\mathcal{L}_1 t} h_{ij}$ whose light cones under H_1 may overlap with the hypercubes. Because each interaction h_{ij} has norm at most $1/\ell_1^\alpha$ by our assumption, the total contribution to the interactions between the cubes x, y is $\mathcal{O}\left(\ell_2^{2d}/\ell_1^\alpha\right)$. Applying the standard finite-range Lieb-Robinson bound on the system of hypercubes, where the maximum energy per interaction is $\mathcal{O}\left(\ell_2^{2d}/\ell_1^\alpha\right)$ and the distance is rescaled by a factor ℓ_2, we obtain the bound for the evolution under $V_{2,I}$:

$$\left\| \mathbb{P}_r e^{\mathcal{L}_I t} |O) \right\| \lesssim \exp\left(\mathcal{O}\left(\frac{\ell_2^{2d}}{\ell_1^\alpha} \right) t - \frac{r}{\ell_2} \right) \equiv e^{\frac{\Delta v t - r}{\ell_2}}, \tag{2.3}$$

where $\Delta v = \mathcal{O}\left(\ell_2^{2d+1}/\ell_1^\alpha\right)$.

After getting the light cone for the evolution under $V_{2,I}$, we now combine it with the evolution under H_1 to obtain the light cone of H_2. Intuitively, the evolutions under H_1 and $V_{2,I}$ for time t may each grow the support radius of an operator by $v_1 t$ and $\Delta v t$ respectively. Therefore, one would expect an operator evolved under H_1 and $V_{2,I}$ consec-

utively, each for time t, may have the support radius at most $(v_1 + \Delta v)t$. In Appendix A, we show that

$$\|\mathbb{P}_r e^{\mathcal{L}_2 t}\,|O)\| = \|\mathbb{P}_r e^{\mathcal{L}_{2,I} t} e^{\mathcal{L}_1 t}\,|O)\| \lesssim e^{\frac{v_2 t - r}{\ell_2}}, \tag{2.4}$$

where

$$v_2 \propto \log(r) v_1 + \Delta v = \log(r) v_1 + \frac{\ell_2^{2d+1}}{\ell_1^{\alpha}} \tag{2.5}$$

and r is the diameter of the lattice Λ. The additional factor of $\log(r)$ (compared to our intuition) comes from the enhancement to the operator spreading due to the increased support size after the first evolution $e^{\mathcal{L}_1 t}$.

Up to this point, we have used the bound Eq. (2.2) for H_1 to prove a bound for H_2 [Eq. (2.4)], which has the same form. Repeating this process, we arrive at similar bounds for H_k $(k = 3, 4, \ldots, n)$:

$$\|\mathbb{P}_r e^{\mathcal{L}_k t}\,|O)\| \lesssim e^{\frac{v_k t - r}{\ell_k}}, \tag{2.6}$$

where the velocity v_k is defined iteratively:

$$v_k \propto \log(r) v_{k-1} + \frac{\ell_k^{2d+1}}{\ell_{k-1}^{\alpha}}. \tag{2.7}$$

Increasing k makes the bound in Eq. (2.6) applicable for longer and longer interactions. However, doing so also increases ℓ_k, resulting in weaker and weaker bounds. In particular,

if $\ell_k > r$, Eq. (2.6) becomes trivial even for $t \leq r/v_k$. Therefore, we stop the iteration at $k = n$ such that ℓ_n is slightly smaller than r. Specifically, we choose n such that $\ell_n = L^n = r/\chi(t,r)$, where $\chi(t,r) > 1$ is a function of t, r. For $v_n t \leq r/2$, the right-hand side of Eq. (2.6) becomes

$$e^{\frac{v_n t - r}{\ell_n}} \lesssim e^{-\frac{r}{2\ell_n}} \lesssim e^{-\frac{1}{2}\chi(t,r)} \lesssim \frac{1}{\chi(t,r)^\omega}, \tag{2.8}$$

where we upper-bound an exponentially decaying function of $\chi(t,r)$ by a power-law decaying function of $\chi(t,r)$ with an exponent $\omega > 0$. Choosing $\chi(t,r) = (r^{\alpha-2d}/t)^\zeta$, where $\zeta > 0$ is an arbitrarily small constant, and $\omega = \frac{\alpha-d}{\zeta(\alpha-2d)}$, we obtain the desired bound

$$\|\mathbb{P}_r e^{\mathcal{L}_n t}|O)\| \lesssim \left(\frac{t}{r^{\alpha-2d}}\right)^{\frac{\alpha-d}{\alpha-2d}}. \tag{2.9}$$

Note that Eq. (2.9) only holds for $t \leq r/2v_n$. To maximize the range of validity of Eq. (2.9), we aim to choose L such that v_n is as small as possible. Without the second term in Eq. (2.7), we would expect v_k to increase by a factor of $\log r$ between iterations. Meanwhile, given $\ell_k = L^k$, the second term in Eq. (2.7) also increases by a factor $L^{2d+1-\alpha}$ in every iteration. Choosing $L^{2d+1-\alpha} \propto \log r$ so that the two terms in Eq. (2.7) have roughly equal contributions to v_k, we expect

$$v_n \propto (\log r)^n \propto L^{n(2d+1-\alpha)} = \left(\frac{r}{\chi(t,r)}\right)^{2d+1-\alpha} \tag{2.10}$$

up to a small logarithmic correction in r. Substituting the earlier choice of $\chi(t,r)$, we

have

$$v_n t \propto r \left(\frac{t}{r^{\alpha-2d}} \right)^{1+o(1)} \leq r, \tag{2.11}$$

where $o(1)$ represents an arbitrarily small constant, for all $t \leq r^{\alpha-2d}$. Therefore, the bound in Eq. (2.9) holds as long as $t \lesssim r^{\alpha-2d}$.

The bound in Eq. (2.9) applies to the Hamiltonian H_n constructed from H by taking interactions of range at most ℓ_n, which is slightly smaller than r for all $t \lesssim r^{\alpha-2d}$. To add interactions of range larger than ℓ_n to the bound, we use the identity [34]:

$$e^{\mathcal{L}t} = e^{\mathcal{L}_n t} + \sum_{i,j:\mathrm{dist}(i,j)>\ell_n} \int_0^t ds\, e^{\mathcal{L}(t-s)} \mathcal{L}_{h_{ij}} e^{\mathcal{L}_n s}, \tag{2.12}$$

where $\mathcal{L}_{h_{ij}} = i[h_{ij}, \cdot]$ is the Liouvillian corresponding to the interaction h_{ij}. We will argue that the contribution from the second term of the right-hand side to the bound on $\|\mathbb{P}_r e^{\mathcal{L}t} |O)\|$ is small.

Note that $\mathcal{L}_{h_{ij}} e^{\mathcal{L}_n s} |O)$ vanishes if $e^{\mathcal{L}_n s} |O)$ has no support on the sites i, j. Suppose site i is closer to the origin than site j. Then, most contributions to the right-hand side of Eq. (2.12) come from terms h_{ij} where i lies within the light cone of $e^{\mathcal{L}_n s} |O)$. Let $\mathcal{V}$ be the volume inside this light cone at time t. Using the triangle inequality on Eq. (2.12), we would arrive at

$$\|\mathbb{P}_r e^{\mathcal{L}t} |O)\| \lesssim \|\mathbb{P}_r e^{\mathcal{L}_n t} |O)\| + \frac{\mathcal{V}t}{\ell_n^{\alpha-d}}, \tag{2.13}$$

where $\mathcal{V}$ is the result of the sum over i inside the light cone, summing over j where

$\mathrm{dist}(i,j) > \ell_n$ gives a factor proportional to $1/\ell_n^{\alpha-d}$, and the integral over time in Eq. (2.12) gives the factor t.

Suppose we can apply the desired light cone $t \gtrsim r^{\alpha-2d}$. Then we can estimate the volume inside the light cone $\mathcal{V} \lesssim t^{d/(\alpha-2d)}$. Substituting it into the above bound together with the value of ℓ_n, we would arrive at

$$\|\mathbb{P}_r e^{\mathcal{L}t}|O)\| \lesssim \left(\frac{t}{r^{\alpha-2d}}\right)^{\frac{\alpha-d}{\alpha-2d}}, \tag{2.14}$$

which gives about the same light cone as in Theorem 1.

However, we are proving Theorem 1 and so cannot yet apply the light cone $t \gtrsim r^{\alpha-2d}$. Instead, we use the light cone from Ref. [19], which is weaker than Theorem 1, to estimate $\mathcal{V}$. Substituting this value of $\mathcal{V}$ into Eq. (2.13), we obtain a *tighter* light cone than that of Ref. [19]. Iteratively using the resulting light cone to estimate $\mathcal{V}$ (see Appendix A for a more detailed derivation), we obtain tighter and tighter bounds. These bounds converge to a stable point that is exactly Eq. (2.14). Therefore, we obtain Theorem 1.

2.3 Discussion

Theorem 1 implies a light cone that can be made arbitrarily close to $t \gtrsim r^{\alpha-2d}$ for all $\alpha \in (2d, 2d+1)$. In addition, Theorem 1 also implies a linear light cone $t \gtrsim r^{1-o(1)}$ for $\alpha \geq 2d+1$, providing an alternative proof to Refs. [22, 23] for two-body Hamiltonians. Together with Refs. [18, 22, 23], we have the final Lieb-Robinson light cone for power-

law interactions:

$$t \gtrsim \begin{cases} \log r & \text{if } d < \alpha \leq 2d \\ r^{\alpha - 2d - o(1)} & \text{if } 2d < \alpha \leq 2d + 1 \\ r & \text{if } \alpha > 2d + 1 \end{cases}, \tag{2.15}$$

which we can saturate, up to subpolynomial corrections, using the protocol for state transfer and entanglement generation in Ref. [26].

Additionally, at any fixed time, our bound decays with distance as $1/r^{\alpha - d - o(1)}$. Because the total strength of the interactions between the origin and all sites that are at distance at least r from the origin already scales as $1/r^{\alpha - d}$, this so-called "tail" of our bound is also optimal.

Our result tightens the constraints on various quantum information tasks in power-law systems, including the growth of connected correlation functions, the generation of topological order, and the digital simulation of local observables. Intuitively, as a local operator evolves, it is mostly constrained to lie within a light cone defined by a Lieb-Robinson bound, with total leakage outside this light cone constrained by the tail of this bound. To simulate the dynamics of such observables, it is sufficient to simulate only the dynamics inside the light cone [21, 25, 35], resulting in a more efficient simulation than simulating the entire lattice. Similarly, the connected correlator between initially local observables remains small during the dynamics if their corresponding light cones have little overlap [25, 29, 36]. Topologically ordered states—those that cannot be distinguished by local observables—would also remain topologically ordered until local observables have

enough time to substantially grow their supports [25, 36]. Crucially, then, Theorem 1, which has a provably optimal light cone and tail, provides the best-known asymptotic constraints for the dynamics of these quantities. The mathematical details of precisely how they are bounded and the improvements that our new bound provides are detailed in Appendix A. Additionally, our result may also provide a tighter constraint on the capacity of quantum communication channels based on power-law interacting spins [37].

While we assume that the Hamiltonian is two-body throughout the chapter, we expect the result extends to general many-body interactions. Specifically, we conjecture that Theorem 1 holds for all Hamiltonians $H = \sum_{X \subset \Lambda} h_X$, where the sum is over all subsets of the lattice and $\sum_{X \ni i,j} \|h_X\| \leq 1/\text{dist}(i,j)^\alpha$ for all $i \neq j$.

Lastly, while Theorem 1 demonstrates the optimality of the single-particle state transfer protocol of [26], other information-theoretic tasks are constrained by tighter light cones. Our techniques may help extend recent progress [25, 38, 39] in constraining the remaining light cone hierarchy that has been demonstrated with power law interactions.

Chapter 3: Saturation of the Lieb-Robinson Bound in Power-Law Interacting Systems

Harnessing entanglement between many particles is key to a quantum advantage in applications including sensing and time-keeping [13, 40], secure communication [41], and quantum computing [42, 43]. For example, encoding quantum information into a multiqubit Greenberger-Horne-Zeilinger-like (GHZ-like) state is particularly desirable as a subroutine in many quantum applications, including metrology [40], quantum computing [44, 45], anonymous quantum communication [46, 47], and quantum secret sharing [48].

The speed at which one can unitarily encode an unknown qubit state $a\,|0\rangle + b\,|1\rangle$ into a GHZ-like state $a\,|00\ldots0\rangle + b\,|11\ldots1\rangle$ of a large system is constrained by Lieb-Robinson bounds [12, 18, 18, 19, 20, 21, 22, 23, 25, 27, 28, 29, 30, 31, 32, 33] and depends on the nature of the interactions in the system. In systems with finite-range interactions and power-law interactions decaying with distance r as $1/r^{\alpha}$ for all $\alpha \geq 2d+1$, where d is the dimension of the system, the Lieb-Robinson bounds imply a linear light cone for the propagation of quantum information [22, 23]. Consequently, in such systems, the linear size of a GHZ-like state that can be prepared from unentangled particles cannot grow faster than linearly with time.

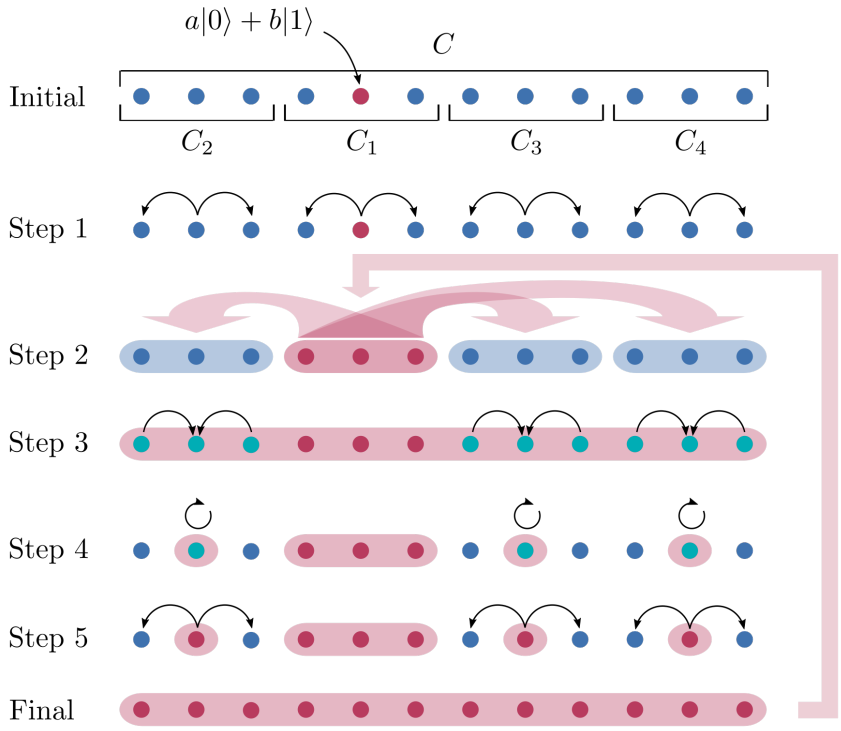

Figure 3.1: A demonstration of our protocol for encoding a qubit into a GHZ-like state in a one-dimensional system C. Initially, the unknown coefficients a, b are encoded in one qubit (red circle) while the other qubits are each initialized in state $|0\rangle$. The first step of the protocol assumes the ability to encode information into GHZ-like states in subsystems $C_1, \ldots, C_4$ using, for example, nearest-neighbor interactions. In step 2, we apply a generalized controlled-PHASE gate [Eq. (3.6)] between the subsystems to "merge" the GHZ-like states into an entangled state between all sites. The last three steps rotate this entangled state into the desired GHZ-like state by concentrating the entanglement in each subsystem onto one qubit, applying single-qubit rotations, and redistributing the entanglement to the rest of the system. Repeatedly feeding the resulting GHZ-like state back into step 2 of the protocol yields larger and larger GHZ-like states.

The Lieb-Robinson bounds become less stringent for longer-range interactions, i.e. those with $\alpha < 2d + 1$. The bounds theoretically allow quantum information to travel a distance r in time t that scales sublinearly with r [18, 19, 21, 29]. However, no protocol in the present literature can saturate these bounds. In particular, existing protocols for $\alpha \in (d, 2d]$ are exponentially slower than what is allowed by the corresponding bounds. Up until now, the existence of this gap between the Lieb-Robinson bounds and the achievable

protocols has meant that at least one of the two is not yet optimal, hinting at either a tighter Lieb-Robinson bound or the possibility of speeding up many quantum information processing tasks.

In this Chapter, we close the gap for all $\alpha \in (d, 2d+1]$ in d dimensions by designing a protocol for encoding an arbitrary qubit into a multiqubit GHZ-like state and, subsequently, transferring information at the limits imposed by the Lieb-Robinson bounds. There are three key implications of the protocol. First, within these regimes of α, it establishes the tightness of the Lieb-Robinson bounds, up to subpolynomial corrections, and effectively puts an end to the fifteen-year search for a tighter bound. Second, our protocol implies optimal designs for future experiments on power-law interacting systems, including trapped ions [49, 50] ($\alpha \in [0, 3]$) in one and two dimensions [1], ultracold atoms in photonic crystals [53, 54], van-der-Waals interacting Rydberg atoms [55, 56] ($\alpha = 6$) in three dimensions [57], as well as the very common case of dipolar interactions in nitrogen-vacancy centers [58], polar molecules [59], and dipole-dipole interacting Rydberg atoms [60] ($\alpha = 3$) in two dimensions. Finally, our protocol implies a lower bound on the gate count in simulating power-law interacting systems on a quantum computer, providing a benchmark for the performance of quantum simulation algorithms.

The structure of the Chapter is as follows. In Sec. 3.1, we define our setting and introduce the main result: the optimal state-transfer time in power-law interacting systems [Eq. (3.3)]. In Sec. 3.2, we describe the corresponding optimal protocol for generating entanglement and subsequently transferring quantum information. At the end of

[1] Quantum computing with trapped ions usually uses resonant addressing and real excitations of the motional modes [51]. On the other hand, for one-dimensional chains of trapped ions, the off-resonant addressing scheme, which results in spin models with tunable approximately-power-law couplings ($\alpha \in [0, 3]$), is popular among recent analog quantum simulation experiments (For example, see Ref. [52]).

Sec. 3.2, we discuss the key ingredients that make the protocol outperform previously known protocols. Readers who are interested in the conceptual implications of the protocol may also skip ahead to Sec. 3.3, where we establish the tightness of existing Lieb-Robinson bounds and discuss implications for other types of speed limits associated with quantum information propagation.

3.1 Optimal State Transfer and Entanglement Generation

We first describe the setting of the problem and the main result in this section. For simplicity, we consider a d-dimensional hypercubic lattice Λ and a two-level system located at every site of the lattice. Our protocol generalizes straightforwardly to all regular lattices. Without loss of generality, we assume that the lattice spacing is one. We consider a power-law interacting Hamiltonian $H(t) = \sum_{i,j\in\Lambda} h_{ij}(t)$, where $h_{ij}(t)$ is a Hamiltonian supported on sites i, j such that, at all times t and for all $i \neq j$, we have $\|h_{ij}\| \leq 1/\text{dist}(i,j)^\alpha$, where $\text{dist}(i,j)$ is the distance between i, j, $\|\cdot\|$ is the operator norm, and $\alpha \geq 0$ is a constant. We use $|\text{GHZ}(a,b)\rangle_S$ to denote the GHZ-like state over sites in $S \subseteq \Lambda$:

$$|\text{GHZ}(a,b)\rangle_S \equiv a\,|\bar{0}\rangle_S + b\,|\bar{1}\rangle_S\,, \tag{3.1}$$

where $|\bar{x}\rangle_S \equiv \bigotimes_{j\in S} |x\rangle_j$ $(x = 0, 1)$ are product states over all sites in S and a, b are complex numbers such that $|a|^2 + |b|^2 = 1$. In particular, we use $|\text{GHZ}\rangle$ to denote the symmetric state $a = b = 1/\sqrt{2}$.

Given a d-dimensional hypercube $C \subseteq \Lambda$ of length $r \geq 1$, we consider the task of

encoding a possibly unknown state $a|0\rangle + b|1\rangle$ of a site $c \in C$ into the GHZ-like state $|\text{GHZ}(a,b)\rangle_C$ over C, assuming that all sites in C, except for c, are initially in the state $|0\rangle$. Specifically, we construct a time-dependent, power-law interacting Hamiltonian $H(t)$ that generates $U(t) = \mathcal{T}\exp\left(-i\int_0^t ds H(s)\right)$ satisfying

$$U(t)\,(a|0\rangle + b|1\rangle)_c\,|\bar{0}\rangle_{C\setminus c} = a|\bar{0}\rangle_C + b|\bar{1}\rangle_C \tag{3.2}$$

at time

$$t(r) \leq K_\alpha \times \begin{cases} \log^{\kappa_\alpha} r & \text{if } d < \alpha < 2d, \\ e^{\gamma\sqrt{\log r}} & \text{if } \alpha = 2d, \text{and} \\ r^{\alpha-2d} & \text{if } 2d < \alpha \leq 2d+1. \end{cases} \tag{3.3}$$

Here, $\gamma = 3\sqrt{d}$, κ_α, and K_α are constants independent of t and r. Additionally, by reversing the unitary in Eq. (3.2) to "concentrate" the information in $|\text{GHZ}(a,b)\rangle$ onto a different site in C, we can transfer a quantum state from $c \in C$ to any other site $c' \in C$ in time $2t$.

3.2 Detailed Steps of the Protocol

The key idea of our protocol (Fig. 3.1) is to recursively build the GHZ-like state in a large hypercube from the GHZ-like states of smaller hypercubes. For the base case, we note that hypercubes of finite lengths, i.e. $r \leq r_0$ for some fixed r_0, can always be generated in times that satisfy Eq. (3.3) for some suitably large (but constant) prefactor

K_α. Assuming that we can encode information into a GHZ-like state in hypercubes of length r_1 in time t_1 satisfying Eq. (3.3), the following subroutine encodes information into a GHZ-like state in an arbitrary hypercube C of length $r = mr_1$ containing c—the site initially holding the phase information a, b. Here m is an α-dependent number to be chosen later.

Step 1: We divide the hypercube C into m^d smaller hypercubes $C_1, \ldots, C_{m^d}$, each of length r_1. Without loss of generality, let C_1 be the hypercube that contains c. Let $V = r_1^d$ be the number of sites in each C_j. In this step, we simultaneously encode a, b into $|\text{GHZ}(a,b)\rangle_{C_1}$ and prepare $|\text{GHZ}\rangle_{C_j}$ for all $j = 2, \ldots, m^d$, which, by our assumption, takes time

$$t_1 \leq K_\alpha \times \begin{cases} \log^{\kappa_\alpha} r_1 & \text{if } d < \alpha < 2d, \\ e^{\gamma\sqrt{\log r_1}} & \text{if } \alpha = 2d, \text{and} \\ r_1^{\alpha-2d} & \text{if } 2d < \alpha \leq 2d+1. \end{cases} \tag{3.4}$$

By the end of this step, the hypercube C is in the state

$$(a\,|\bar{0}\rangle + b\,|\bar{1}\rangle)_{C_1} \bigotimes_{j=2}^{m^d} \frac{|\bar{0}\rangle_{C_j} + |\bar{1}\rangle_{C_j}}{\sqrt{2}}. \tag{3.5}$$

Step 2: Next, we apply the following Hamiltonian to the hypercube C:

$$H_2 = \frac{1}{(mr_1\sqrt{d})^\alpha} \sum_{j=2}^{m^d} \sum_{\mu \in C_1} \sum_{\nu \in C_j} |1\rangle\langle 1|_\mu \otimes |1\rangle\langle 1|_\nu . \tag{3.6}$$

This Hamiltonian effectively generates the so-called controlled-PHASE gate between the

hypercubes, with C_1 being the control hypercube and $C_2, \ldots, C_{m^d}$ being the target hypercubes. We choose the interactions between qubits in Eq. (3.6) to be identical for simplicity. If the interactions were to vary between qubits, we would simply turn off the interaction between C_1 and C_j once the total phase accumulated by C_j reaches π [2]. The prefactor $1/(mr_1\sqrt{d})^\alpha$ ensures that this Hamiltonian satisfies the condition of a power-law interacting Hamiltonian. It is straightforward to verify that, under this evolution, the state of the hypercube C rotates to

$$a\,|\bar{0}\rangle_{C_1} \bigotimes_{j=2}^{m^d} \frac{|\bar{0}\rangle_{C_j} + |\bar{1}\rangle_{C_j}}{\sqrt{2}} + b\,|\bar{1}\rangle_{C_1} \bigotimes_{j=2}^{m^d} \frac{|\bar{0}\rangle_{C_j} - |\bar{1}\rangle_{C_j}}{\sqrt{2}} \tag{3.7}$$

after time $t_2 = \pi d^{\alpha/2}(mr_1)^\alpha/V^2$.

The role of power-law interactions in our protocol can be inferred from the value of t_2. Intuitively, the speed of simultaneously entangling hypercube C_1 with hypercubes $C_2, \ldots, C_{m^d}$ is enhanced by the $V^2 = r_1^{2d}$ couplings between the hypercubes. However, the strength of each coupling, proportional to $1/(mr_1)^\alpha$, is suppressed by the maximum distance between the sites in C_1 and those in $C_2, \ldots, C_{m^d}$. With a small enough α, the enhancement due to V^2 overcomes the suppression of power-law interactions, resulting in a small entanglement time t_2. In particular, when $\alpha < 2d$, t_2 actually *decreases* with r_1, implying that Step 2 would be faster in later iterations if we were to keep m constant.

[2] Because only the total accumulated phase matters in choosing the evolution time, we also expect the protocol to be robust against experimental errors such as uncertainties in the positions of individual particles: If the position of each particle is known up to a precision $\varepsilon \ll 1$, the total worst-case error in the accumulated phase scales as $t(r_1^{2d}/r^\alpha) \times (\varepsilon/r)$, with r_1 being the length of each hypercubes and r being the minimum distance between them. The result is a relative phase error proportional to ε/r, which becomes smaller and smaller as the distance between the hypercubes increases. Moreover, we expect the relative error to be even smaller in the commonly occurring situation when uncertainties in the positions are uncorrelated between different particles

To obtain the desired state $|\text{GHZ}(a,b)\rangle_C$, it remains to apply a Hadamard gate on the effective qubit $\{|\bar{0}\rangle_{C_j}, |\bar{1}\rangle_{C_j}\}$ for $j = 2, \ldots, m^d$. We do this in the following three steps by first concentrating the information stored in hypercube C_j onto a single site $c_j \in C_j$ (Step 3), then applying a Hadamard gate on c_j (Step 4), and then unfolding the information back onto the full hypercube C_j (Step 5).

Step 3: By our assumption, for each hypercube C_j ($j = 2, \ldots, m^d$) and given a designated site $c_j \in C_j$, there exists a (time-dependent) Hamiltonian H_j that generates a unitary U_j such that

$$(\psi_0 |0\rangle + \psi_1 |1\rangle)_{c_j} |\bar{0}\rangle_{C_j \setminus c_j} \xrightarrow{U_j} \psi_0 |\bar{0}\rangle_{C_j} + \psi_1 |\bar{1}\rangle_{C_j} \tag{3.8}$$

for all complex coefficients ψ_0 and ψ_1, in time t_1 satisfying Eq. (3.4). By linearity, this property applies even if C_j is entangled with other hypercubes. Consequently, backward time evolution under H_j generates $U_j^\dagger$, which "undoes" the GHZ-like state of the jth hypercube:

$$\psi_0 |\bar{0}\rangle_{C_j} + \psi_1 |\bar{1}\rangle_{C_j} \xrightarrow{U_j^\dagger} (\psi_0 |0\rangle + \psi_1 |1\rangle)_{c_j} |\bar{0}\rangle_{C_j \setminus c_j} \tag{3.9}$$

for any ψ_0, ψ_1. In this step, we simultaneously apply $U_j^\dagger$ to C_j for all $j = 2, \ldots, m^d$. These unitaries rotate the state of C to

$$a |\bar{0}\rangle_{C_1} \bigotimes_{j=2}^{m^d} |+\rangle_{c_j} |\bar{0}\rangle_{C_j \setminus c_j} + b |\bar{1}\rangle_{C_1} \bigotimes_{j=2}^{m^d} |-\rangle_{c_j} |\bar{0}\rangle_{C_j \setminus c_j}, \tag{3.10}$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$.

Step 4: We then apply a Hadamard gate, i.e.

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}, \tag{3.11}$$

to the site c_j of each hypercubes C_j, $j = 2, \ldots, m^d$. These Hadamard gates can be implemented arbitrarily fast since we do not assume any constraints on the single-site terms of the Hamiltonian. The state of C by the end of this step is

$$a\left|\bar{0}\right\rangle_{C_1}\bigotimes_{j=2}^{m^d}\left|0\right\rangle_{c_j}\left|\bar{0}\right\rangle_{C_j\setminus c_j}+b\left|\bar{1}\right\rangle_{C_1}\bigotimes_{j=2}^{m^d}\left|1\right\rangle_{c_j}\left|\bar{0}\right\rangle_{C_j\setminus c_j}. \tag{3.12}$$

Step 5: Finally, we apply U_j again to each hypercube C_j ($j = 2, \ldots, m^d$) to obtain the desired GHZ-like state:

$$a\left|\bar{0}\right\rangle_{C_1}\bigotimes_{j=2}^{m^d}\left|\bar{0}\right\rangle_{C_j}+b\left|\bar{1}\right\rangle_{C_1}\bigotimes_{j=2}^{m^d}\left|\bar{1}\right\rangle_{C_j}=\left|\mathrm{GHZ}(a,b)\right\rangle_{C}. \tag{3.13}$$

At the end of this routine, we have implemented the unitary satisfying Eq. (3.2) in time

$$t = 3t_1 + t_2 = 3t_1 + \pi d^{\alpha/2} m^{\alpha} r_1^{\alpha-2d}. \tag{3.14}$$

We now consider three cases corresponding to different ranges of α and show that if $t_1(r_1)$ satisfies Eq. (3.3), then $t(r)$ also satisfies Eq. (3.3).

For $\alpha \in (2d, 2d+1]$, we have $t_1 \leq K_\alpha r_1^{\alpha-2d}$. Choosing $m > 1$ to be a constant

integer, we have

$$t \leq \left(\frac{3K_\alpha}{m^{\alpha-2d}} + \pi d^{\alpha/2} m^{2d} \right) (mr_1)^{\alpha-2d} \leq K_\alpha r^{\alpha-2d}, \tag{3.15}$$

where we require $m > 3^{1/(\alpha-2d)}$ and choose

$$K_\alpha \geq \frac{\pi d^{\alpha/2} m^{2d}}{1 - \frac{3}{m^{\alpha-2d}}} = \frac{\pi d^{\alpha/2} m^{\alpha}}{m^{\alpha-2d} - 3}. \tag{3.16}$$

For $\alpha \in (d, 2d)$, we choose m to scale with r_1 such that $r_1^{\lambda-1} < m \leq 2r_1^{\lambda-1}$ where $\lambda = 2d/\alpha$. The length of the larger cube C is then $r = mr_1 > r_1^\lambda$ and, therefore, the total time is

$$t \leq 3K_\alpha \log^{\kappa_\alpha} r_1 + \pi(2\sqrt{d})^\alpha r_1^{(\lambda-1)\alpha+\alpha-2d} \tag{3.17}$$

$$\leq \frac{4K_\alpha}{\lambda^{\kappa_\alpha}} \log^{\kappa_\alpha} \left(r_1^\lambda\right) \leq K_\alpha \log^{\kappa_\alpha} r, \tag{3.18}$$

where we choose $\kappa_\alpha = \log 4 / \log(2d/\alpha)$ and assume $K_\alpha \log^{\kappa_\alpha} r_1 \geq \pi(2\sqrt{d})^\alpha$ to simplify the expression. We note that the factor $\log 4$ in the definition of κ_α can be made arbitrarily close to $\log 3$ by increasing K_α.

Finally, for $\alpha = 2d$, we choose m such that $\exp(\frac{\gamma}{2d}\sqrt{\log r_1}) \leq m \leq 2\exp(\frac{\gamma}{2d}\sqrt{\log r_1})$, where $\gamma = 3\sqrt{d}$. Substituting $t_1 \leq K_\alpha \exp(\gamma\sqrt{\log r_1})$ into Eq. (3.14), we have

$$t \leq \left(3K_\alpha + 2^\alpha \pi d^{\alpha/2}\right) e^{\gamma\sqrt{\log r_1}}. \tag{3.19}$$

Assuming $r_1 \geq \exp(8/d)$, it is straightforward to prove that $\gamma\sqrt{\log r_1} \leq \gamma\sqrt{\log(mr_1)} - 2$.

Applying this condition on the above inequality, we have

$$t \leq \frac{1}{e^2}\left(3K_\alpha + 2^\alpha \pi d^{\alpha/2}\right) e^{\gamma\sqrt{\log r}} \leq K_\alpha e^{\gamma\sqrt{\log r}}, \tag{3.20}$$

where $r = mr_1$ is the length of the resulting GHZ-like state and we chose $K_\alpha \geq 2^\alpha \pi d^{\alpha/2}/(e^2 - 3)$. Equations (3.15), (3.18) and (3.20) prove that t satisfies Eq. (3.3). Repeatedly applying this routine yields larger and larger GHZ-like states.

Before discussing the implications of our protocol, we would like to explain intuitively the main sources of its improvement relative to existing protocols. In our protocol, we simultaneously encode the information into the GHZ-like state over C_1 and create the symmetric GHZ states over other multiqubit subsystems $C_2, \ldots, C_{m^d}$. As a result, the implementation of the controlled operations in step 2 (Fig. 3.1) is enhanced quadratically by the volume of each subsystems. In contrast, the protocol in Ref. [24] applies controlled operations between a large subsystem and individual remaining sites of the system, resulting in the implementation time scaling only linearly with the volume of the subsystem.

On the other hand, while the state transfer protocol in Refs. [23, 25] also applies controlled operations between large subsystems and is, therefore, sped up quadratically by the subsystem volume, it only uses qubits in small neighborhoods around the source and the target of the transfer. In our protocol, we maximize the size of the resulting GHZ-like state at the end of each iteration by allowing m to depend on α and on the size of the existing GHZ-like states. When we use the protocol for state transfer, this strategy results in most of the qubits between the source and the target sites participating in the transfer,

Tasks	Known light cones	Previous best protocols	Our protocol
Encoding into a GHZ-like state	$t \gtrsim \begin{cases} \log r & \alpha \in (d, 2d] \text{ [18]} \\ r^{\alpha-2d} & \alpha \in (2d, 2d+1) \text{ [22, 61]} \end{cases}$	$t \sim \begin{cases} r^{\alpha-d} & \alpha \in (d, d+1) \text{ [24]} \\ r & \alpha \in [d+1, 2d+1) \end{cases}$	$t \sim \begin{cases} \text{polylog}(r) & \alpha \in (d, 2d) \\ e^{\gamma\sqrt{\log r}} & \alpha = 2d \\ r^{\alpha-2d} & \alpha \in (2d, 2d+1) \end{cases}$
Preparing a known GHZ-like state	$t \gtrsim \begin{cases} \log r & \alpha \in (d, 2d] \text{ [18, 36]} \\ r^{\alpha-2d} & \alpha \in (2d, 2d+1) \text{ [36, 61]} \end{cases}$	Same as encoding into a GHZ-like state	Same as above
State transfer	Same as encoding into a GHZ-like state	$t \sim \begin{cases} r^{\frac{\alpha(\alpha-d)}{\alpha+d}} & \alpha \in (d, d+1) \text{ [25]} \\ r^{\frac{\alpha}{2d+1}} & \alpha \in [d+1, 2d+1) \text{ [23, 25]} \end{cases}$	Same as above
State transfer (no initialization)	$t \gtrsim \begin{cases} r^{\frac{2\alpha-2d}{2\alpha-d+1}} & \alpha \in (d, 2d] \text{ [38]} \\ r^{\alpha-2d} & \alpha \in (2d, 2d+1) \text{ [21]} \\ r^{\alpha-1} & \alpha \in (1, 2], d = 1 \text{ [62]} \\ r & \alpha \in (2, 3), d = 1 \text{ [62]} \end{cases}$	$t \sim r \ \ \forall \alpha \in (d, 2d+1)$	Not applicable

Table 3.1: A summary of known bounds and protocols in the regime $\alpha \in (d, 2d+1)$ for several information-propagation tasks: encoding an unknown qubit state into a GHZ-like state (row 1), preparing a known GHZ-like state (row 2), state transfer assuming we can initialize intermediate qubits (row 3), and state transfer given intermediate qubits in arbitrary states (i.e. so-called universal state transfer [25], row 4). The tasks of encoding information into GHZ-like states and quantum state transfer with initialization are constrained by the Lieb-Robinson bounds. Preparing a known GHZ-like state, being potentially easier than encoding unknown information into GHZ-like states, is—at least at present—sometimes bounded by a weaker light cone [21, 29]. On the other hand, state transfer given intermediate qubits in arbitrary states (i.e. universal state transfer) is more difficult than state transfer with initialized intermediate qubits and is bounded by the more stringent Frobenius light cone [25]. The bounds on encoding information into GHZ-like states (except Ref. [22]) also apply to general k-body interactions. All listed bounds also hold not just for qubits, but for all finite-level systems. For $d < \alpha \leq 2d+1$, our protocol saturates (up to subpolynomial corrections) the known bounds, thus proving the optimality of both the protocol and the bounds.

significantly speeding up the protocol.

3.3 Discussion

We now discuss the performance and the implications of our protocol (summarized in Table 3.1). First, our protocol allows for encoding an unknown qubit into a multiqubit GHZ-like state and, subsequently, performing state transfer at unprecedented speeds. For $d < \alpha < 2d$, which applies, for example, to dipole-dipole interactions ($\alpha = 3$) in two dimensions and to the effective interactions between trapped ions ($\alpha \in [0, 3]$) in one

and two dimensions, our protocol encodes information into GHZ-like states and transfers information in polylogarithmic time, exponentially faster than protocols available in the literature. Even for the seemingly weakly long-range interactions with $\alpha = 2d$, such as van der Waals interactions between Rydberg atoms ($\alpha = 6$) in three dimensions, our protocol still takes only subpolynomial time to entangle an entire system and to transfer a quantum state. When applied to the preparation of GHZ states, these speedups enable potential improvements to quantum sensors built from nitrogen-vacancy centers [63, 64], Rydberg atoms [65, 66], and polar molecules [67], as well as to atomic clocks based on trapped ions [68].

Optimal quantum information processing.—The optimality of our protocol for $\alpha \in (d, 2d+1)$ also lays the foundation for optimal quantum information processing in power-law interacting systems [49, 50]. Using quantum state transfer between auxiliary qubits and encoding qubits into large GHZ-like states as subroutines, our protocol leads to optimal implementations of quantum gates between distant qubits in large quantum computers. In particular, the faster encoding of information into a GHZ-like state of ancillary qubits speeds up [45] the implementation of the quantum fanout—a powerful multiqubit quantum gate [69]. At the same time, the faster state transfer speeds up [24] the constructions of multiscale entanglement renormalization ansatz (MERA) states, commonly used to represent highly entangled—including topologically ordered [70]—states [71, 72, 73]. Specifically, we can implement a fanout gate [45] on qubits in a hypercube of volume n and prepare a MERA state [24] on these qubits in time $t \sim \text{polylog}(n)$ for $\alpha \in (d, 2d)$, $t \sim e^{\frac{\gamma}{\sqrt{d}}\sqrt{\log n}}$ for $\alpha = 2d$—which are both exponential speedups compared to the previous best—and $t \sim n^{(\alpha-2d)/d}$ for $\alpha \in (2d, 2d+1)$. The optimality of these operations is again

guaranteed (up to subpolynomial corrections) by the matching lower limits imposed by the Lieb-Robinson bounds [24, 45].

In practice, using single-site Hamiltonians to implement the echoing technique of Ref. [24], the controlled-PHASE gate in step 2 of our protocol can be realized starting from time-independent power-law interactions between all sites of the system. The protocol therefore does not require explicit time-dependent control of individual two-qubit Hamiltonians, making it appealing for implementation on available experimental platforms. However, because the diameter of the GHZ-like state increases by more than twofold in every iteration of the protocol, the scaling in Eq. (3.3) may only be observed in large systems.

Information-propagation speed limits.— Conceptually, since our protocol saturates (up to subpolynomial corrections) the Lieb-Robinson bounds for $d < \alpha \leq 2d + 1$ for all d, we demonstrate, for the first time, the tightness of these fundamental bounds in these regimes. In particular, the subpolynomial entanglement time for $\alpha \leq 2d$ disproves the conjecture in Refs. [74, 75], where a gap in the understanding of the heating times and the effective generators of dynamics in periodically driven, power-law interacting systems had suggested the existence of a tighter Lieb-Robinson bound with an algebraic light cone in this regime of α. We discuss in more detail below what could have resulted in this gap in our understanding.

Since the best known generalizations of these bounds to k-body, power-law interacting Hamiltonians—those described by $H = \sum_X h_X$, where the sum is over all subsets $X \subset \Lambda$ of at most k sites and $\sum_{X \ni i,j} \|h_X\| \leq 1/\mathrm{dist}(i,j)^\alpha$ for all $i \neq j$—have the same scaling as the best known 2-body bounds when $d < \alpha \leq 2d$ [18] (see also Table 3.1),

the scaling of our 2-body protocol is also optimal even if one allows for k-body interactions. In other words, in this regime of α, allowing for k-body interactions cannot enable a qualitative speedup relative to 2-body interactions.

Our protocol also generalizes straightforwardly from two-level to arbitrary finite-level systems. Given a q-level system at each site of the lattice, we can unitarily encode an arbitrary state $|\psi\rangle_c = \sum_{\ell=0}^{q-1} a_\ell |\ell\rangle$ of site $c \in C$, where a_ℓ are complex coefficients and C is a hypercube of linear size r, into a multi-qudit state

$$|\psi\rangle_c |\bar{0}\rangle_{C\setminus c} \to \sum_{\ell=0}^{q-1} a_\ell \left|\bar{\ell}\right\rangle_C \tag{3.21}$$

in time $t(r)$ satisfying Eq. (3.3). This can be done by replacing the Hamiltonian in Eq. (3.6) with

$$\frac{1}{(mr_1\sqrt{d})^\alpha} \sum_{j=2}^{m^d} \sum_{\mu\in C_1} \sum_{\nu\in C_j} \sum_{\ell,\ell'=0}^{q-1} \ell\ell' |\ell\rangle \langle\ell|_\mu \otimes |\ell'\rangle \langle\ell'|_\nu \tag{3.22}$$

and replacing the single-qubit Hadamard gate in step 4 by a q-by-q discrete Fourier transform matrix. Since the Lieb-Robinson bounds have the same light cones for any finite-level systems, our protocol also saturates these bounds for $\alpha \in (d, 2d]$ in d dimensions.

In our protocol, we assume that $a|0\rangle + b|1\rangle$ is a possibly unknown state. Encoding such a state into the GHZ-like state is at least as hard as generating a GHZ-like state with known coefficients a, b. In fact, the latter task is not known to be sufficient for state transfer and, therefore, is not *directly* constrained by the Lieb-Robinson bounds. Instead, one often indirectly obtains a speed limit for this task by applying the Lieb-Robinson

bounds on the growth of two-point connected correlators [18, 25, 36]. Our protocol for encoding into a GHZ-like state saturates (up to subpolynomial corrections) the bound $t \gtrsim \log r$ [18, 36] on the growth of connected correlators when $d < \alpha \leq 2d$, implying that knowing the coefficients a, b does not speed up the preparation of the GHZ-like state in this regime.

We also note that our protocol violates the so-called Frobenius light cone, initially derived in Ref. [25] for $\alpha > 3/2$ in one dimension as part of a hierarchy of speed limits for different types of information propagation in long-range interacting systems and later extended to regimes of smaller α in Ref. [38]. The Frobenius bound, which considers information propagation from the operator-spreading perspective, constrains information-propagation tasks that are more demanding than the tasks that saturate the Lieb-Robinson bound, and therefore has a more stringent light cone. For example, quantum state transfer given intermediate qubits in arbitrary initial states (i.e. universal state transfer) is constrained by the Frobenius light cone, whereas state transfer assuming initialized intermediate qubits is constrained by the Lieb-Robinson bound and can actually violate the Frobenius light cone [25] (see also Table 3.1). Determining which of the bounds tightly constrains a given task is still an active area of research. The protocol in this manuscript proves for the first time that the task of encoding information into GHZ-like state—which is at least as hard as state transfer with initialization—is not constrained by the Frobenius light cone, but is instead tightly constrained (up to subpolynomial corrections) by the Lieb-Robinson bound. In particular, when $d < \alpha < 2d$, our protocol proves that state transfer with initialization can be implemented exponentially faster than state transfer without initialization, which is constrained by polynomial light cones in this

regime [25, 38]. Furthermore, since our protocol for encoding into a GHZ-like state can also be used to prepare a known GHZ-like state, our protocol also proves for the first time that preparing a known GHZ-like state is not constrained by the Frobenius light cone.

Resource lower bound for quantum simulation.—Our protocol also gives the first known example of a lower bound on the gate count for simulating power-law systems on a quantum computer: it takes $\Omega(n)$ elementary quantum gates to simulate an n-qubit power-law system evolving for time $t \geq t_*$, where

$$t_* = \begin{cases} \Theta\left(\log^{\kappa_\alpha} n\right) & \text{if } d < \alpha < 2d, \\ \Theta\left(e^{\gamma\sqrt{(\log n)/d}}\right) & \text{if } \alpha = 2d, \text{ and} \\ \Theta\left(n^{\alpha/d-2}\right) & \text{if } 2d < \alpha \leq 2d+1, \end{cases} \tag{3.23}$$

to constant error. Indeed, if an algorithm could use fewer than $\Omega(n)$ quantum gates to perform the simulation for times within $t = t_*$ satisfying Eq. (3.23), we could use the algorithm to simulate our protocol and prepare an n-qubit GHZ state. However, since an n-qubit GHZ state must take $\Omega(n)$ quantum gates to prepare, we would arrive at a contradiction.

Lower bounds on the simulation gate count are valuable benchmarks for the performance of quantum algorithms. Ref. [76] gives an algorithm for simulating the time evolution of finite-range interacting Hamiltonians, the gate count of which was shown to be optimal via a matching lower bound. To date, despite progressively more efficient quantum simulation algorithms [21, 77] in recent literature, no saturable lower bounds are known for power-law systems. For example, the analysis of the Suzuki-Trotter product

formulas in Ref. [77] results in upper bounds

$$g_\alpha = \begin{cases} \mathcal{O}\left(n^{2+o(1)}t^{1+o(1)}\right) & \text{if } d < \alpha \leq 2d, \\ \mathcal{O}\left((nt)^{1+d/(\alpha-d)+o(1)}\right) & \text{if } \alpha > 2d, \end{cases} \tag{3.24}$$

for simulating an n-qubit power-law system for time t. At $t = t_*$ given in Eq. (3.23), the corresponding upper bounds reduce to

$$g_\alpha = \begin{cases} \mathcal{O}\left(n^{2+o(1)}\right) & \text{if } d < \alpha \leq 2d, \\ \mathcal{O}\left(n^{\alpha/d+o(1)}\right) & \text{if } 2d < \alpha \leq 2d+1. \end{cases} \tag{3.25}$$

The gap between this state-of-the-art upper bound and our lower bound $\Omega(n)$ hints at the possibility of a more efficient algorithm for simulating power-law systems.

Chapter 4: A Hierarchy of Speed Limits for the Propagation of Quantum Information

Until recently, it was unknown whether or not there existed a critical value of the power-law exponent α above which a linear light cone is present. Hastings and Koma [18] first demonstrated a light cone whose velocity diverges exponentially in distance for α greater than the lattice dimension, d. Progressive improvements yielded a series of algebraic light cones for $\alpha > 2d$, which tend to a linear light cone in the limit as $\alpha \to \infty$ [19, 21]. After numerical simulations suggested the existence of a sharp linear light cone [22, 78, 79], a proof of generic linear light cones was found for systems with interaction exponent $\alpha > 2d + 1$ [22, 23].

Complementary to the Lieb-Robinson bounds are protocols that achieve the (asymptotically) fastest allowable rates of quantum information processing. One such dynamical task is quantum state transfer, which has been used experimentally to demonstrate the transmission of entanglement in quantum systems [80]. These protocols can be directly connected to the Lieb-Robinson bound [24, 81] and have been a standard way to benchmark the sharpness of these bounds.

The goal of this Chapter is to answer an important questions: do the tightest light-cone bounds imply correspondingly tight bounds on interesting measures of informa-

tion spreading, such as quantum state transfer or scrambling? In other words, are Lieb-Robinson bounds optimal in practice for constraining quantum information dynamics?

Surprisingly, the answer to this question is "no." In this Chapter, we show that quantum information can spread at arbitrarily large "velocities" once the power law exponent $\alpha < 2d+1$, thus proving the tightness of the recent bounds [22, 23]. We also show that a *Frobenius* bound can give tighter constraints on quantum state transfer tasks—as well as many-body quantum chaos—than *Lieb-Robinson* bounds. We prove that the light cone given by the Frobenius bound is linear for $\alpha > \frac{5}{2}$ in $d = 1$, and conjecture the generalization $\alpha > \frac{3}{2}d + 1$ for higher dimensions. Additionally, in systems that are described by non-interacting bosons or fermions, we prove a linear *free-particle* light cone for $\alpha > d + 1$. All of these cutoffs in this hierarchy of linear light cones are tight: see Figure 4.1.

These results immediately demonstrate that the long-observed mismatch between Lieb-Robinson bounds and state-transfer protocols that aim to saturate the bounds, such as that of Ref. [24], is not entirely a limitation of our creativity or mathematical prowess, but is rather linked to a fundamental property of nature. There are, simply put, *multiple notions of locality* in systems with long-range interactions. Furthermore, the tensions among these localities manifest themselves within a range of α that is easily accessible in experiment. This unexpected result is the key finding of this chapter [25].

The hierarchy of linear light cones we demonstrate is not only a profound property of nature, but also has important applications for quantum technologies. For example, systems with long-range interactions can be hard problems to simulate, both on classical and quantum computers. Proving the tightness of the linear Lieb-Robinson light cone at

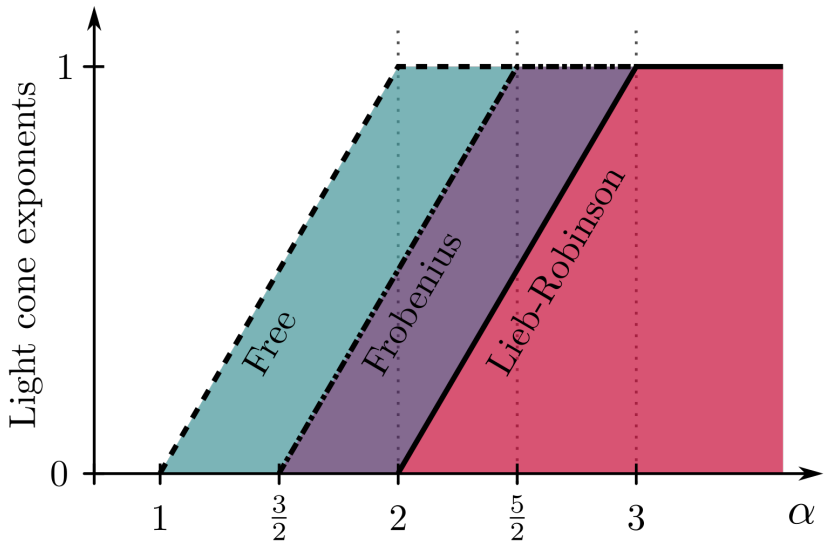

Figure 4.1: The hierarchy of linear light cones in one dimension; we say that a light cone has exponent γ if $\|[A_0(t), B_r]\|$ is large only when $t \gtrsim r^\gamma$. The plot depicts the exponents of the Lieb-Robinson light cone (solid line) [22], the Frobenius light cone from Theorem 4 (dot-dashed line), and the free light cone from Theorem 5 (dashed line) as functions of α in one dimension. The free light cone is known to be a tight bound for all α. We also show that the Lieb-Robinson and Frobenius light cones are not linear below $\alpha = 3$ and $\alpha = \frac{5}{2}$ respectively.

$\alpha > 2d + 1$ proves that a two-dimensional gas of atoms with dipole-dipole interactions can never be simulated as easily as one with local interactions with a provably small error. At the same time, the hierarchy of light cones reveals that *some* problems are much easier to simulate than had previously been realized. As a specific example, the Bose-Hubbard model (with long-range hopping) has been argued to be so difficult that its efficient simulation would serve as a demonstration of quantum supremacy [82]. Our light cones show that it is not difficult to simulate the low-density Bose-Hubbard model for $\alpha > d$, whereas previously this was only known for $\alpha > 2d$ [83]; as a result, we have substantially constrained the parameter space in which quantum supremacy can be demonstrated. This result constrains how and when atoms with dipole-dipole interactions

trapped in a two-dimensional optical lattice can perform hard quantum computation or simulation.

High fidelity quantum state transfer can be used to build fast remote quantum gates, which can significantly speed up a large-scale quantum computer. There is a growing interest in designing fully-connected quantum computers that take advantage of long-range interactions among physical qubits [15], and finding the optimal quantum state-transfer protocols using long-range interactions is a crucial part of the design. The hierarchy of light cones we find reveals the fundamental inadequacy of Lieb-Robinson bounds for constraining *universal* state transfer algorithms, which transfer the state of a single qubit independently of the states of other qubits. We develop a quantum walk formalism for constraining universal state-transfer protocols, and obtain parametrically better bounds than the Lieb-Robinson bound. Furthermore, the framework that has been initiated in this work also reveals novel state-transfer protocols with desirable properties. Specifically, we present a new method for using long-range interactions for state transfer that has two experimentally desirable features. Firstly, our new protocol takes place in a constrained subspace of a many-body Hilbert space that is naturally realized in atomic platforms with a conserved magnetization. Secondly, the protocol is extraordinarily robust to perturbations in the Hamiltonian, a desirable feature on account of the low-precision tunable couplers present in near-term quantum information processors.

Platforms with long-range interactions have been proposed as natural quantum simulators which approximately realize $\alpha = 0$ (i.e. all-to-all) interactions. Systems with such a complete breakdown of locality can be highly desirable. For example, they may simulate quantum gravity via the holographic correspondence [84] and may enable the

production of metrologically useful entanglement via spin-squeezing [13, 14, 85, 86, 87]. An important open question is how small α needs to be for locality to break down to a degree sufficient for realizing a particular application or particular physics. For example, are dipolar $1/r^3$ interactions in a given 1D, 2D, or 3D system sufficiently non-local? Our results indicate that the answer to these questions may depend on whether there are additional constraints in the system. Indeed, in a highly constrained subspace at high total spin in an SU(2)-symmetric model, we expect that the constraints arising from locality are stronger than the Lieb-Robinson light cone suggests, similar to the stronger light cone that arises for non-interacting particles. Therefore, in such constrained models, reaching non-locality may require a lower value of α compared to unconstrained models.

Lastly, we emphasize that given that there is a hierarchy of different notions of locality, exquisite care must be taken to analyze and interpret experimental results in long-range interacting quantum systems.

4.1 An Introduction to the Hierarchy

We now provide a heuristic overview of why the hierarchy of light cones arises, along with a myriad of additional applications of these results in near-term quantum simulation experiments. The remainder of the chapter then contains the rigorous proof of all results, along with a brief conclusion.

For illustrative purposes, let us first consider a one-dimensional spin-$\frac{1}{2}$ chain with two-body long-range interactions. Such models naturally arise in experiments, for example, using the nuclear spin-$\frac{1}{2}$ of an appropriate atom. Letting $X_i^\beta := (X_i, Y_i, Z_i)$ denote

the three Pauli matrices acting on the spin on the i^{th} site, we can consider a very broad family of time-dependent Hamiltonians of the form

$$H(t) = \sum_{i,\beta} h_i^\beta(t) X_i^\beta + \sum_{i,j,\beta,\gamma} \frac{J_{ij}^{\beta\gamma}(t)}{|i-j|^\alpha} X_i^\beta X_j^\gamma. \tag{4.1}$$

Roughly speaking, if the coefficients $J_{ij}^{\beta\gamma}$ are all of the same order, we call this Hamiltonian a model with long-range interactions of power law exponent α. Remarkably, even though every spin is coupled with every other spin, this model is, for many practical purposes, *local* for α sufficiently large [22, 23] (indeed, this Hamiltonian even becomes finite-range in the limit $\alpha \to \infty$). But, what do we mean by locality? And how small can α get before locality breaks down? We will see that, in fact, there are multiple notions of locality, depending on the specific quantities of interest: different information processing tasks are sped up by long range interactions at different values of α.

The Lieb-Robinson light cone.— A sensible notion of locality is to demand that any local perturbation, acting at site x will only influence physics at sites within distance vt of the original site x, after an amount of time t [29]. This notion of locality is imposed by the original Lieb-Robinson bound [12], which implies a "linear light cone": the quantity $\|[X_0(t), X_r]\|$ is small for $r > vt$, where X_0 and X_r are local operators on lattice sites 0 and r respectively and $\|\cdot\|$ denotes the operator norm (the largest magnitude of the operator's eigenvalues). More precisely, a linear light cone here means that for any small (but finite) value of ε, we can find a finite velocity v such that $\|[X_0(t), X_r]\| < \varepsilon$ for $|t| < r/v$.

As noted previously, recent works have established linear Lieb-Robinson light cones

for $\alpha > 2d+1$ [22, 23]. The first main result of this work is to prove that linear Lieb-Robinson light cones are *only* guaranteed for any $\alpha > 2d+1$. We prove this by explicitly constructing a Hamiltonian $H(t)$—of the form (4.1), generalized to any d—such that

$$\|[X_0(t), X_r]\| \gtrsim \frac{t^{2d+1}}{r^\alpha}. \tag{4.2}$$

for two sites separated by a distance r (Theorem 3).

The construction of $H(t)$ that achieves Eq. (4.2) can be broken down into three steps. In the first step, we use time $\mathcal{O}(1)$ to expand the operator X_0 to an operator A_1 supported on $O(t^d)$ sites located in a ball $\mathcal{B}_1$ of radius t (Fig. 4.2). We then push this operator into another ball $\mathcal{B}_2$ of radius $\mathcal{O}(t)$, which is centered around site X_r, using the Hamiltonian

$$H_2(t) = \sum_{i\in\mathcal{B}_1}\sum_{j\in\mathcal{B}_2} \frac{Z_i Z_j}{|i-j|^\alpha}. \tag{4.3}$$

Finally, we contract the operator onto site X_r in time $O(t)$. By a direct calculation, we show that, at the end of this process and to the lowest order in t, $\|[[X_0(t), X_r]]\|$ is proportional to the nested commutator $t\|[A_2, [H_2(t), A_1]]\|$, where

$$A_{1,2} = \prod_{i\in\mathcal{B}_{1,2}} X_i. \tag{4.4}$$

The nested commutator can be bounded:

$$\|[A_2, [H(t), A_1]]\| \sim t^d \left\| \left[A_2, \sum_{j\in\mathcal{B}_2} \frac{Z_j}{r^\alpha} \right] \right\| \sim \sum_{j\in\mathcal{B}_2} \frac{t^{2d}}{r^\alpha}, \tag{4.5}$$

and hence we obtain Eq. (4.2). We conclude that in a time t, it is faster to use long-range interactions than it is to use finite-range ones to grow $\|[X_0(t), X_r]\|$ when $\alpha < 2d + 1$.

In Sec. 6.1, we show that connected correlation functions of the form

$$C(t) = \langle\psi(t)| X_0 X_r |\psi(t)\rangle - \langle\psi(t)| X_0 |\psi(t)\rangle \langle\psi(t)| X_r |\psi(t)\rangle \gtrsim \frac{t^{2d+1}}{r^\alpha} \tag{4.6}$$

can be achieved, even when the initial state $|\psi(0)\rangle$ does not have any entanglement between sites 0 and r. Because such correlation functions are routinely measured in quantum simulation experiments, this result resolves a long-standing issue of when the nonlinear light cones for correlations can occur with long-range interactions. For example, the experiment in Ref. [88] suggests that $\alpha \approx 1$ marked the transition between linear and nonlinear light cones for spin correlations in a 11-site long-range Ising model. Our result implies that other quantum systems with long-range interactions can transmit information *much faster* than this Ising model.

Another important application of Lieb-Robinson bounds is to design efficient approximation algorithms for simulating quantum many-body dynamics, with either a classical computer [89] or a quantum computer [76]. Given an initial state $|\psi\rangle$ and Hamiltonian H of the form (4.1), we consider the task of estimating the expectation value of the time-evolved observable $\langle A(t)\rangle := \langle\psi| U(t)^\dagger A U(t) |\psi\rangle$ on a quantum computer, where $U(t) = \mathrm{e}^{-\mathrm{i}Ht}$ is the time translation operator generated by H (assuming it does not depend on time). When A is a local operator, Lieb-Robinson bounds suggest that $\langle A(t)\rangle$ should depend only on the "local" information stored in the wave function: one may as well trace out and ignore the sites sufficiently far away. If a Lieb-Robinson bound implies

we can trace out all sites a distance $> vt$ away from the support of A, computation of $\langle A(t)\rangle$ requires a small fraction of the resources needed to construct the full $U(t)$ acting on the entire many-body Hilbert space. Proposition 3 makes this intuition precise and constrains the computational resources needed for faithful quantum simulation.

In Sec. 6.7, we use Lieb-Robinson bounds to constrain the minimum time τ^* it takes to create topologically ordered states from topologically trivial ones. This result is of great practical value for experiments either studying topological matter or building topological quantum memories and topological quantum computers. In finite-range interacting systems, τ^* scales linearly with the system size [36]. We extend this result in Proposition 4 to power-law interacting systems with $\alpha > 3d + 1$.

In Sec. 6.8, we bound the spatial decay of correlation functions in a ground state of a gapped quantum phase with long-range interactions. In Ref. [18], the authors show that in a time-independent power-law Hamiltonian with an exponent α and a spectral gap between the ground state and the first excited state, the correlations between distant sites in the ground state of the system also decay with the distance as a power law, with an exponent lower bounded by $\alpha' < \alpha$. Yet no experiment and no numerical calculation has found a gapped system demonstrating correlation decay with exponent strictly less than α. We prove that it is indeed impossible to saturate this bound; we show that the correlation exponent is lower bounded by $\alpha' = \alpha$ whenever $\alpha > 2d$.

The Frobenius light cone.— In Sec. 4.4, we turn to a stronger notion of light cone, inspired by recent developments in the theory of many-body quantum chaos [90, 91].

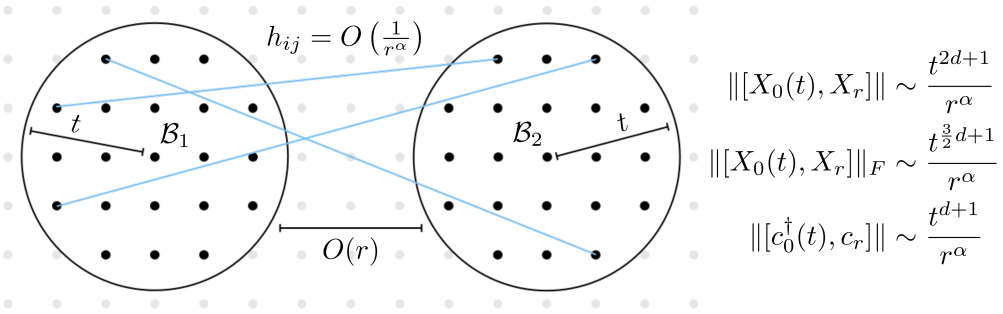

Figure 4.2: The norm of the interaction between two balls $\mathcal{B}_{1,2}$ of size t, separated by a distance r, determine the shape of the light cone. The critical values of α after which this norm becomes large differs depending on whether we use the operator or Frobenius norm, and whether the system is interacting or free.

Instead of the operator norm, we consider the Frobenius norm

$$\|[X_0(t), X_r]\|_{\mathrm{F}} := \sqrt{\frac{\mathrm{tr}([X_0(t), X_r]^\dagger [X_0(t), X_r])}{\mathrm{tr}(\mathbb{I})}}. \tag{4.7}$$

This Frobenius norm, normalized by dimension, can be interpreted as the out-of-time-ordered correlation (OTOC) function used to probe early time chaos in many-body systems [90, 91] or, equivalently, as the "fraction" of the operator $X_0(t)$ that has support on the site r. More intuitively, this OTOC can be understood by the following thought experiment: consider an initial quantum state $|\psi\rangle$, and perturb this quantum state by two operators: first, the local operator X_r (which flips spin r in the conventional Z-basis), and then the Heisenberg-evolved operator X_0, which amounts to flipping spin 0 at a later time t. Does the order of these operations matter? Clearly not if $t = 0$: $X_0 X_r = X_r X_0$. However, if the operations occur at different times t, the effect could be significant: $X_0(t)X_r|\psi\rangle$ might be a very different quantum state than $X_r X_0(t)|\psi\rangle$. We can quantify how far apart these two states are in Hilbert space by asking for the typical length of $[X_0(t), X_r]|\psi\rangle$,

or the value of $C = \langle\psi|[X_0(t), X_r]^\dagger[X_0(t), X_r]|\psi\rangle$. A suitable notion of "typical" is to choose a random initial state in the Hilbert space. Averaging over all initial conditions amounts to replacing $\mathrm{tr}(|\psi\rangle\langle\psi| \cdots) \to \mathrm{tr}(\frac{1}{\dim(\mathcal{H})} \cdots)$. Hence the average value of C is given by (4.7). Mathematically, the Frobenius norm gives the average of the squared eigenvalues, while the operator norm used in Lieb-Robinson bounds is the maximal eigenvalue. Certainly, the Frobenius norm is always smaller: $\|[X_0(t), X_r]\|_{\mathrm{F}} \leq \|[X_0(t), X_r]\|$.

Remarkably, in long-range interacting systems, we can show that the Frobenius norm is not only smaller by a constant prefactor, but is rather constrained by *parametrically stronger* bounds. Indeed, we prove in Section 4.4 that $\|[X_0(t), X_r]\|_{\mathrm{F}}$ is bounded inside of an even stricter light cone, which is linear in one-dimensional models with two-body interactions so long as $\alpha > \frac{5}{2}$. When X_0 is replaced by an operator on infinitely many sites $0, -1, -2, \ldots$, we also demonstrate the optimality of this bound, up to subalgebraic corrections.

To understand how the Frobenius light cone is deduced, let us revisit the argument for the Lieb-Robinson light cone. Modifying Eq. (4.5) to use the Frobenius norm, we observe that

$$t\|[A_2, [H_2(t), A_1]]\|_{\mathrm{F}} \sim t \left\| \left[A_2, \sum_{j \in \mathcal{B}_2} \frac{t^d Z_j}{r^\alpha} \right] \right\|_{\mathrm{F}} \sim t \sqrt{\sum_{j \in \mathcal{B}_2} \left(\frac{t^d}{r^\alpha} \right)^2} \sim \frac{t^{\frac{3}{2}d+1}}{r^\alpha}. \tag{4.8}$$

Hence, in d dimensions, the Frobenius norm of the operator grows faster than in local models using long-range interactions once $\alpha < \frac{3}{2}d + 1$. For $d = 1$, Theorem 4 proves that this intuition is correct: the Frobenius light cone is linear when $\alpha > \frac{5}{2}$. In between $\frac{3}{2} < \alpha < \frac{5}{2}$, in $d = 1$, this theorem also guarantees that the Frobenius light cone expands

no faster than $t \sim r^{\alpha-\frac{3}{2}}$ (up to logarithmic corrections).

The mathematical method used to prove the Frobenius light cone is based on an interpretation of the time evolution equation for operators as a many-body quantum walk governing the time evolution of a probability distribution. By bounding the growth of expectation values in this probability distribution using techniques from classical probability theory, we constrain the growth of Eq. (4.7). This represents a radical shift in perspective compared with the conventional Lieb-Robinson theorem, which is based on applying the triangle inequality in an appropriate interaction picture (see e.g. Refs. [19, 23]).

Since the Frobenius norm (squared) gives infinite temperature OTOCs, the Lieb-Robinson light cone *is not relevant* for infinite temperature many-body quantum chaos and the growth of operators. A careful determination of bounds on quantum chaos and operator spreading is essential for building on recent experimental progress in measuring OTOCs [92, 93] and quantum information scrambling [17] to design optimal information scramblers. Such work will be crucial in developing quantum simulators of holographic quantum gravity [84].

As emphasized before, many quantum state transfer tasks, including a "background-independent" state transfer where $X_i(t) = X_j$, $Y_i(t) = Y_j$ and $Z_i(t) = Z_j$ (hence state i is transferred to j independently of all other qubits), are constrained by the Frobenius light cone, which is tighter than the Lieb-Robinson light cone: see Theorem 4.

The Free light cone.— Finally, we consider the light cone in systems of non-interacting particles. While these systems are rich enough such that they sometimes lie beyond the regime of computability for classical computers, their dynamics can be essentially reduced to the motion of a single particle. Returning to the same setup of Figure 4.2, we

may again estimate when the linear light cone fails by computing the weight of a single particle hopping the distance $\sim r$ from the ball $\mathcal{B}_1$ to $\mathcal{B}_2$ after time t:

$$t\left\| \sum_{i\in\mathcal{B}_1, j\in\mathcal{B}_2} \frac{c_j^\dagger c_i}{r^\alpha} \right\| \lesssim \frac{t^{d+1}}{r^\alpha}. \tag{4.9}$$

Here $c_i^\dagger$ and c_i are the creation and the annihilation operators for the non-interacting particles. Following our previous logic, the free particle is constrained within a linear light cone when $\alpha > d+1$. We rigorously prove that the free light cone is linear for $\alpha > d+1$ in Theorem 5, and prove that no linear light cones exist for $\alpha < d+1$ in Theorem 6. When combined, these two theorems also prove that for $d < \alpha < d+1$, the form of the light cone is no worse than $t \sim r^{\alpha-d}$, and that no further improvement on the exponent $\alpha - d$ can be found.

Specifically, in Theorem 6 (Section 4.5.3), we show that this estimated growth rate is achieved by a novel quantum state-transfer protocol involving a single particle. The protocol works by successively spreading a particle to larger and larger regions of the lattice, each time doubling the number of sites sharing the particle (Fig. 4.3). Specifically, after the kth step of the protocol at time t_k, an operator $c_0^\dagger$ originally supported at the origin would become

$$c_0^\dagger(t_k) \propto \sum_{\text{sites } x \text{ in a cube of length } O(2^k)} c_x^\dagger, \tag{4.10}$$

where the precise set of sites x is depicted in Fig. 4.3. After spreading the particle to a square large enough to cover both the origin and the target site, we simply reverse the protocol to concentrate the particle on the target site. In each step of the protocol,

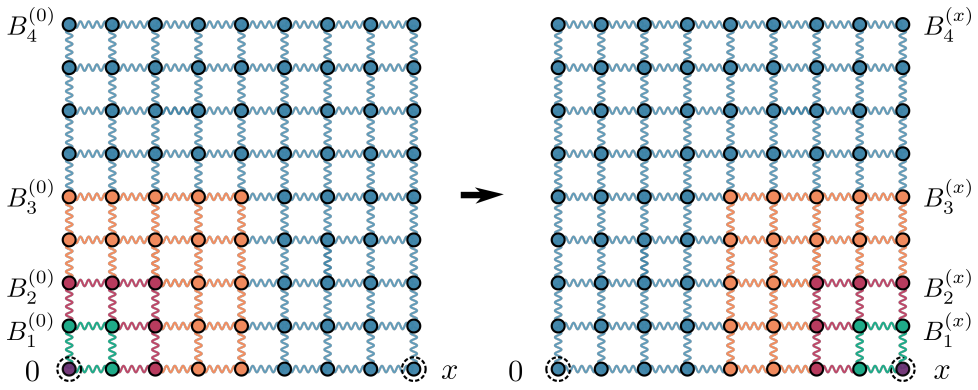

Figure 4.3: An illustration of our single-particle state-transfer protocol in $d = 2$ dimensions. Through 4 steps, we redistribute a particle initially at 0 to the square $B_1^{(0)}$, then to $B_2^{(0)}, B_3^{(0)}$, and finally to $B_4^{(0)}$, each time doubling the size of the region sharing the particle. We use different colors to mark the additional sites that the particle spreads to in different steps. We then reverse the process to concentrate the particle on the target site x and thereby achieve a perfect state transfer. The protocol is enhanced by the volume r^d of the squares, with r being a typical size of the squares. This enhancement offsets the penalty $1/r^\alpha$ due to the power-law constraint, resulting in a superlinear state-transfer protocol when $\alpha < d + 1$.

the weaker interactions due to the power-law constraint are well compensated by the volume of the squares, making the protocol superlinear for all $\alpha < d + 1$. As emphasized in the introduction, this state-transfer protocol has (at least) two appealing features for experimental implementation, and could enhance the performance of quantum computing architectures assisted by long-range interactions [15].

The free light cone is also relevant for early time dynamics in low density models of interacting fermions or bosons, which are readily realized in experiment. The observed slowdown of dynamics between the Lieb-Robinson and free light cones makes the Hubbard model exponentially easier to simulate in experimentally relevant regimes (e.g. polar molecules) at early times, with implications for demonstrating quantum supremacy (Section 6.6).

4.2 Formal Preliminaries

We now more carefully introduce the problem that we address in this chapter. First, we will give a precise definition of a many-body quantum system with long-range interactions. We need to first define the distance between two points. Formally, we do so as follows. Let Λ be the vertex set of a d-dimensional lattice graph with edge set E_Λ. A lattice graph (Λ, E_Λ) is a graph which is invariant under d-dimensional discrete translations: mathematically speaking, $\mathbb{Z}^d \subseteq \mathrm{Aut}(\Lambda, E_\Lambda)$, where Aut denotes the group of graph isomorphisms from (Λ, E_Λ) to itself. We assume that all vertices have finite degree in E_Λ, and that $|\Lambda/\mathbb{Z}^d| < \infty$, i.e. the unit cell has a finite number of vertices, and every vertex has a finite number of (nearest) neighbors. This graph imbues a natural notion of distance, which we will use for the rest of the chapter. Let $\mathcal{D} : \Lambda \times \Lambda \to \mathbb{Z}^+$ denote the shortest path length between two vertices, also known as the Manhattan metric.

A many-body quantum system is then defined by placing a finite-dimensional quantum system (e.g. a qubit) on every vertex in Λ. Formally we define a many-body Hilbert space

$$\mathcal{H} := \bigotimes_{i \in \Lambda} \mathcal{H}_i, \tag{4.11}$$

where we assume that $\dim(\mathcal{H}_i) < \infty$. In this chapter, we will be especially interested in the dynamics of the operators acting on $\mathcal{H}$. Let $\mathcal{B}$ denote the set of all Hermitian operators acting on $\mathcal{H}$. $\mathcal{B}$ is a real vector space, and we denote operators $\mathcal{O} \in \mathcal{B}$ with $|\mathcal{O})$ whenever we wish to emphasize that they should be thought of as vectors. A basis for $\mathcal{B}$ can be found as follows: let T_i^a denote the generators of $\mathrm{U}(\dim(\mathcal{H}_i))$ where $a = 0$ denotes

the identity operator, which gives a complete basis for Hermitian operators on the local Hilbert space $\mathcal{H}_i$. $\mathcal{B}$ is simply the tensor product of all these local bases of Hermitian operators:

$$\mathcal{B} = \text{span}\left\{\bigotimes_{i\in\Lambda} T_i^{a_i},\ \text{for all } \{a_i\}\right\}. \tag{4.12}$$

For subset $X \subset \Lambda$, we define $\mathcal{B}_X$ to be the set of all basis vectors which act non-trivially only on the sites of X:

$$\mathcal{B}_X := \text{span}\left\{\bigotimes_{i\in X} T_i^{a_i},\ \text{for all } \{a_i \neq 0\}\right\}. \tag{4.13}$$

We define the projectors [34]

$$\mathbb{P}_i|\otimes T_k^{a_k}) := \begin{cases} |\otimes T_k^{a_k}) & a_i \neq 0 \\ 0 & a_i = 0 \end{cases}, \tag{4.14}$$

which return the part of the operator that acts non-trivially on site i:

$$\mathbb{P}_i\mathcal{O} = \mathcal{O} - \frac{1}{\dim(\mathcal{H}_i)}\underset{i}{\text{tr}}\ \mathcal{O}. \tag{4.15}$$

For a general subset $X \subset \Lambda$, the projectors

$$\mathbb{P}_X := \sum_{Y\in\mathbb{Z}_2^X : |Y|>0} (-1)^{|Y|+1}\prod_{j\in Y}\mathbb{P}_j \tag{4.16}$$

act similarly, and return the part of the operator which acts non-trivially on the subset X.

It was proven in Ref. [22] that when $|X| < \infty$,

$$\|\mathbb{P}_X \mathcal{O}\|_\infty \leq 2\|\mathcal{O}\|_\infty, \tag{4.17}$$

where $\|\cdot\|_\infty$ is again the operator norm. We will often drop the ∞ subscript for convenience. In addition, we can relate the commutator in the Lieb-Robinson bound to the projection of an operator using the identity

$$\|[\mathcal{O}_X, \mathcal{O}_Y]\| \leq 2\|O_X\|\|\mathbb{P}_X \mathcal{O}_Y\|, \tag{4.18}$$

which holds for all operators $\mathcal{O}_X \in \mathcal{B}_X, \mathcal{O}_Y \in \mathcal{B}_Y$.

We define the Hamiltonian $H : \mathbb{R} \to \mathcal{B}$ as

$$H(t) := \sum_{X \subset \Lambda} H_X(t), \tag{4.19}$$

where $H_X(t) : \mathbb{R} \to \mathcal{B}_X$. $H(t)$ is said to be q-local if $H_X(t) = 0$ for all $|X| > q$: physically speaking, the Hamiltonian operator contains at most q-body interactions. The Hamiltonian generates time evolution on $\mathcal{B}$ according to the Heisenberg equation of motion for operators: we define the Liouvillian $\mathcal{L}(t)$ as the generator of time evolution,

$$\mathcal{L}(t)|\mathcal{O}) := |\mathrm{i}[H(t), \mathcal{O}]). \tag{4.20}$$

We define the time evolved operator $\mathcal{O}(t) : \mathbb{R} \to \mathcal{B}$ as the solution to the differential

equation

$$\frac{\mathrm{d}\mathcal{O}(t)}{\mathrm{d}t} := \mathcal{L}(t)\mathcal{O}(t), \quad \mathcal{O}(0) := \mathcal{O}. \tag{4.21}$$

We say that the Hamiltonian H has long-range interactions with exponent α if

$$\alpha = \sup \Bigg\{ \alpha_0 \in (0,\infty) : \text{there exists } 0 < h < \infty \text{ such that} \\ \sum_{X:\{i,j\}\subseteq X} \|H_X(t)\| \leq \frac{h}{\mathcal{D}(i,j)^{\alpha_0}}, \text{ for all } t \in \mathbb{R} \Bigg\}, \tag{4.22}$$

where $\mathcal{D}(i,j)$ denotes the distance between i,j. In physics we often say that the interaction has exponent α when, assuming only two-body interactions, $H_{\{i,j\}} \lesssim h\mathcal{D}(i,j)^{-\alpha}$; strictly speaking though, any Hamiltonian with exponent α_2, according to this loose definition, also has exponent $\alpha_1 < \alpha_2$. The formal definition Eq. (4.22) avoids this unwanted feature and assigns a unique exponent α to every problem.

The following identities, which we state without proof, will be useful in the discussion that follows:

Proposition 1 (Sums over power laws [18, 21])**.** *If $\alpha > d$, for any Λ and $\mathcal{D}$, there exist $0 < C_1, C_2 < \infty$ such that:*

$$\sum_{j\in\Lambda:\mathcal{D}(i,j)>r} \frac{1}{\mathcal{D}(i,j)^{\alpha}} < \frac{C_1}{r^{\alpha-d}}, \tag{4.23}$$

$$\sum_{k\in\Lambda\setminus\{i,j\}} \frac{1}{\mathcal{D}(i,k)^{\alpha}\mathcal{D}(j,k)^{\alpha}} < \frac{C_2}{\mathcal{D}(i,j)^{\alpha}}. \tag{4.24}$$

4.3 The Lieb-Robinson Light Cone

We begin by discussing the strictest light cone on the commutators of local operators, representing the generalization of the Lieb-Robinson theorem [12] to systems with long-range interactions.

4.3.1 The Linear Light Cone

The following proposition controls the growth of commutator norms in a Hamiltonian system with long-range interactions.

Proposition 2. *Let $X, Y \subset \Lambda$ be disjoint with $\mathcal{D}(X, Y) := r$; $\mathcal{O}_X$ be an operator supported on X obeying $\|\mathcal{O}_X\| = 1$; $\mathcal{O}_X(t)$ be the time-evolved version of $\mathcal{O}_X$ under a power-law Hamiltonian with an exponent $\alpha > 2d+1$. There exist constants $0 < \bar{v}, c < \infty$ such that, for time evolutions generated by Eq. (4.20) obeying Eq. (4.22),*

$$\|\mathbb{P}_Y\left(\mathcal{O}_X(t)\right)\| \leq c|X|\frac{t^{d+1}\log^{2d} r}{(r-\bar{v}t)^{\alpha-d}}. \tag{4.25}$$

Proof. We begin by recalling the following theorem (recast in the language of projectors):

Theorem 2 (Linear light cone [23])**.** *Eq. (4.25) holds for a single-site operator, i.e. when $|X| = 1$.*

While the proof presented in Ref. [23] applied only to time-independent Hamiltonians, the generalization to time-dependent models is immediate from their results. Next, we show the following general result.

Lemma 1. *If for all* $x \in X$, $\|\mathbb{P}_Y|\mathcal{O}_x(t))\| \leq f(t, \mathcal{D}(x, Y))$, *then there exist* $0 < K < \infty$ *such that*

$$\|\mathbb{P}_Y|\mathcal{O}_X(t))\| \leq K \sum_{x\in X} f(t, \mathcal{D}(x, Y)). \tag{4.26}$$

Proof. For pedagogical reasons, we demonstrate the proof on a system of spin-$1/2$ particles with $K = 9/2$. However, the proof applies to any system with finite local Hilbert space dimensions [22]. Let $\{S_j : j = 1, \ldots, d_Y - 1\}$ denote the $d_Y - 1 = 4^{|Y|} - 1$ nontrivial Pauli strings supported on Y. Then [22]

$$\begin{aligned}\|\mathbb{P}_Y\,|\mathcal{O}_X(t))\| &= \|\frac{1}{2d_Y}\sum_{j=1}^{d_Y-1}[S_j,[S_j,\mathcal{O}_X(t)]]\| \leq \frac{1}{2d_Y}\sum_{j=1}^{d_Y-1} 2\underbrace{\|S_j\|}_{=1}\|[S_j,\mathcal{O}_X(t)]\| \\ &\leq \frac{1}{d_Y}\sum_{j=1}^{d_Y-1}\|[\mathcal{O}_X, S_j(-t)]\| \leq \frac{2}{d_Y}\sum_{j=1}^{d_Y-1}\|\mathbb{P}_X\,|S_j(-t))\|. \end{aligned}\tag{4.27}$$

Next, we shall prove that

$$\|\mathbb{P}_X\,|S_j(-t))\| \leq 3\sum_x \|\mathbb{P}_x\,|S_j(-t))\|. \tag{4.28}$$

To do so, we assign an (arbitrary) ordering of the sites in X: i.e. if $X = \{x_1, \ldots, x_n\}$, we choose $x_1 < x_2 < \cdots < x_n$. Let $\tilde{X}_x = \{x' \in X : x' > x\}$ be a subset of X consisting of sites in X that are greater than x. We rewrite

$$\mathbb{P}_X = \sum_x (1 - \mathbb{P}_{\tilde{X}_x})\mathbb{P}_x, \tag{4.29}$$

and therefore we have

$$\|\mathbb{P}_X\left|S_j(-t)\right)\| \leq \sum_x \|(1-\mathbb{P}_{\tilde{X}_x})\mathbb{P}_x\left|S_j(-t)\right)\| \leq \sum_x 3\|\mathbb{P}_x\left|S_j(-t)\right)\|. \tag{4.30}$$

In the last line, we have used that $\|\mathbb{P}_X\mathcal{O}\| \leq 2\|\mathcal{O}\|$ whenever $|X| < \infty$ [22], along with the triangle inequality. Plugging this back into the earlier equation, we have

$$\begin{aligned}\|\mathbb{P}_Y\left|\mathcal{O}_X(t)\right)\| &\leq \frac{6}{d_Y}\sum_{j=1}^{d_Y-1}\sum_x \|\mathbb{P}_x\left|S_j(-t)\right)\| \leq \frac{6}{d_Y}\sum_{j=1}^{d_Y-1}\sum_x \frac{1}{8}\sum_{P_x}\|[P_x,[P_x,S_j(-t)]]\| \\ &\leq \frac{3}{2}\underbrace{\frac{1}{d_Y}\sum_{j=1}^{d_Y-1}}_{\leq 1}\sum_x\sum_{P_x}\|\mathbb{P}_{S_j}P_x(t)\| \leq \frac{9}{2}\sum_x f(t,\text{dist}\left(()\,x,Y\right)),\end{aligned} \tag{4.31}$$

where $P_x \in \{X_x, Y_x, Z_x\}$ denotes one of the three Pauli matrices on site x. In the second from the last line, we have used the assumption $\|\mathbb{P}_{S_j}P_x(t)\| \leq f(t,\text{dist}\left(()\,x,Y\right))$. □

Combining Theorem 2 with Lemma 1 proves Eq. (4.25), which is tighter than a result of Ref. [23] when applied to general operators that are supported on many sites. □

4.3.2 A Fast Operator-Spreading Protocol

Proposition 2 proves that the support of an operator $\mathcal{O}_i(t)$ is only large inside of a linear light cone when $\alpha > 2d+1$. Our first main result is the following theorem, which proves the optimality (up to subalgebraic corrections) of that result.

Theorem 3. *Let* $\dim(\mathcal{H}_i) = 2$ *for all* $i \in \Lambda$*, and let* X_0 *and* X_r *be two Pauli-*X *operators supported on two sites* i *and* j *respectively, obeying* $\mathcal{D}(i,j) = r$*. For all* $\alpha > d$*, there exists a time-dependent Hamiltonian* $H(t)$ *obeying Eq. (4.22) and constants* $0 < K, K' <$

∞ *such that for* $3 < t < K' r^{\alpha/(1+2d)}$,

$$\|[X_0(t), X_r]\| \geq K \frac{t^{1+2d}}{r^\alpha}. \tag{4.32}$$

Proof. We prove the theorem by constructing a fast operator-spreading protocol, which follows three steps, as depicted in Figure 4.4. In each step, we evolve the operator using a power-law Hamiltonian for time $t/3$. For simplicity, we assume $t/3 := \ell \in \mathbb{Z}^+$, and assume that $\ell < \frac{1}{2}r$.

Step 1—. In time $t/3$, we use a unitary $\mathcal{U}_1$ to spread the operator X_0 to $\prod_{i\in\mathcal{B}_\ell} X_i$, where $\mathcal{B}_\ell$ is a ball of radius ℓ centered at site 0. We denote the volume of this ball by $\mathcal{V} := |\mathcal{B}_\ell|$. The unitary $\mathcal{U}_1$ can be implemented using a series of controlled-NOT operators (CNOT) among nearest neighbors in the lattice. Note that a CNOT gate $U_{\mathrm{CNOT},i,j}$ for neighbors i and j acts as follows:

$$U^\dagger_{\mathrm{CNOT},i,j} X_i U_{\mathrm{CNOT},i,j} := X_i X_j. \tag{4.33}$$

Under the conditions of Eq. (4.22), this CNOT gate can be implemented in a time step of $O(1)$.

Step 2—. In the next $t/3$ interval, we apply $\mathcal{U}_2 = \prod_{j\in\mathcal{B}_\ell} U_j(\tau)$ on the operator, where

$$U_j(\tau) := \cos(\tau\Theta) + \mathrm{i}\sin(\tau\Theta) Z_j, \tag{4.34a}$$

$$\Theta := \sum_{k\in\tilde{\mathcal{B}}_\ell} Z_k, \tag{4.34b}$$

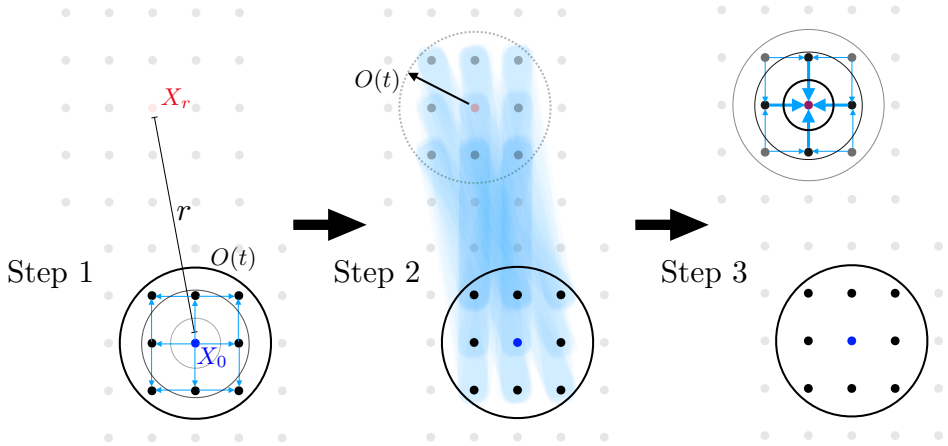

Figure 4.4: A protocol for rapid growth of the commutator norm using two-body long-range interactions. Step 1: we use CNOT gates between nearest neighbor sites to spread a single Pauli X_0 to a Pauli string $XX\cdots X$ supported on every site inside a ball of radius $O(t)$ centered at X_0. Step 2: we use pairwise ZZ interactions between all sites in the two balls, located distance $O(r)$ apart, which adds an operator of norm $\sim O(t^{2d+1}/r^{\alpha})$ into the second ball a distance r away. Step 3: we invert Step 1 in the outer ball, pushing all of the operator weight in the outer ball onto a single site.

$\tilde{\mathcal{B}}_\ell$ is another ball of radius d centered around the site at distance r, and

$$\tau = \frac{t}{3(2r)^{\alpha}}. \tag{4.35}$$

It is straightforward to verify that $U_j(\tau)$ is a unitary, since

$$U_j(\tau) = \exp\left[-\mathrm{i}\tau \sum_{y\in\tilde{\mathcal{B}}_\ell} Z_j Z_y\right]. \tag{4.36}$$

Since $Z_j Z_y$ commutes with $Z_{j'} Z_{y'}$ for all $j, j' \in \mathcal{B}_\ell$ and $y, y' \in \tilde{\mathcal{B}}_\ell$, $U_j(\tau)$ and $U_{j'}(\tau)$ can be implemented simultaneously. In other words, the unitary $\mathcal{U}_2$ can be generated by a power-law Hamiltonian within time $t/3$: the factor of $2r$ in Eq. (4.35) is present because the maximal distance between two sites in $\mathcal{B}_\ell$ and $\tilde{\mathcal{B}}_\ell$ is $r + 2\ell < 2r$. The evolved version

of the operator under this unitary is

$$\mathcal{U}_2^\dagger \left(\prod_{j \in \mathcal{B}_\ell} X_j \right) \mathcal{U}_2 = \prod_{j \in \mathcal{B}_\ell} \left[\cos(2\tau\Theta) X_j + \sin(2\tau\Theta) Y_j \right]. \tag{4.37}$$

Step 3—. In the final $t/3$, we apply a unitary $\mathcal{U}_3$ which is the inverse of $\mathcal{U}_1$, up to its action on $\tilde{\mathcal{B}}_\ell$ instead of $\mathcal{B}_\ell$. It is easier to instead think of evolving the final operator X_r under $\mathcal{U}_3^{-1}$, which does not change the commutator norm $\|[X_0(t), X_r]\|$. Therefore, after time t, we get the commutator norm:

$$\|[X_1(t), X_r]\| = \| \left[\mathcal{U}_2^\dagger \mathcal{U}_1^\dagger X_0 \mathcal{U}_1 \mathcal{U}_2, \mathcal{U}_3 X_r \mathcal{U}_3^\dagger \right] \| \tag{4.38}$$

$$= \Bigg\| \underbrace{\left[\prod_{j \in \mathcal{B}_\ell} \left[\cos(2\tau\Theta) X_j + \sin(2\tau\Theta) Y_j \right], \prod_{k \in \tilde{\mathcal{B}}_\ell} X_k \right]}_{\equiv C} \Bigg\|. \tag{4.39}$$

To lower bound the norm of C, we consider the matrix elements of C in the eigenbasis of Pauli Z operators. We observe that $\langle e| C |00\ldots0\rangle = 0$ for all computational basis states $|e\rangle$ of the two balls except for $|e\rangle = |11\ldots1\rangle$. Hence,

$$\begin{aligned} \langle 11\ldots1| C |00\ldots0\rangle &= \left[\cos(2\tau\mathcal{V}) - i \sin(2\tau\mathcal{V}) \right]^{\mathcal{V}} - \left[\cos(2\tau\mathcal{V}) + i \sin(2\tau\mathcal{V}) \right]^{\mathcal{V}} \\ &= -2 \sum_{k \text{ odd}}^{\mathcal{V}} \binom{\mathcal{V}}{k} i^k \sin(2\tau\mathcal{V})^k \cos(2\tau\mathcal{V})^{\mathcal{V}-k} := a. \end{aligned} \tag{4.40}$$

Therefore, C is block diagonal and has eigenvalues $\pm|a|$ in the sector $\{|00\ldots0\rangle\}, \{|11\ldots1\rangle\}$. We note that to the lowest-order, $|a| \approx \mathcal{V}^2\tau \propto t^{2d+1}/r^\alpha$. Therefore, this operator-spreading protocol saturates the Lieb-Robinson bound in Proposition 2.

To make the statement rigorous, we lower bound the norm of C:

$$\begin{aligned}\|C\| \geq |a| &\geq 2\mathcal{V}\sin(2\tau\mathcal{V})\cos(2\tau\mathcal{V})^{\mathcal{V}-1} - \sum_{k \text{ odd}, k\geq 3}^{\mathcal{V}} \binom{\mathcal{V}}{k} \sin(2\tau\mathcal{V})^k \cos(2\tau\mathcal{V})^{\mathcal{V}-k} \\ &\geq 2\mathcal{V}\sin(2\tau\mathcal{V})\cos(2\tau\mathcal{V})^{\mathcal{V}-1} - \frac{\mathcal{V}^3}{6}\sin(2\tau\mathcal{V})^3 \sum_{k \text{ even}}^{\mathcal{V}-3} \binom{\mathcal{V}-3}{k} \sin(2\tau\mathcal{V})^k \cos(2\tau\mathcal{V})^{\mathcal{V}-3-k} \\ &\geq 2\mathcal{V}\sin(2\tau\mathcal{V})\cos(2\tau\mathcal{V})^{\mathcal{V}-1} - \frac{\mathcal{V}^3}{6}\sin(2\tau\mathcal{V})^3[\sin(2\tau\mathcal{V}) + \cos(2\tau\mathcal{V})]^{\mathcal{V}-3}. \end{aligned} \tag{4.41}$$

Now we require that $\mathcal{V}^2\tau = \varepsilon < 1/2$, which is equivalent to $t^{2d+1} \lesssim r^\alpha$. Under this condition,

$$\cos(2\tau\mathcal{V})^{\mathcal{V}-1} \geq \left(1-\tau^2\mathcal{V}^2\right)^{\mathcal{V}} = \left(1-\frac{\varepsilon^2}{\mathcal{V}^2}\right)^{\mathcal{V}} \geq \frac{1}{2}, \qquad \text{(for all } \mathcal{V} \geq 1 > \varepsilon^2/10), \tag{4.42a}$$

$$[\sin(2\tau\mathcal{V}) + \cos(2\tau\mathcal{V})]^{\mathcal{V}-3} \leq (1+2\tau\mathcal{V})^{\mathcal{V}} \leq \left(1+\frac{2\varepsilon}{\mathcal{V}}\right)^{\mathcal{V}} \leq \mathrm{e}^{2\varepsilon}, \tag{4.42b}$$

$$\tau\mathcal{V} \leq \sin(2\tau\mathcal{V}) \leq 2\tau\mathcal{V}. \tag{4.42c}$$

Therefore,

$$\begin{aligned}\|C\| &\geq 2\mathcal{V}(\tau\mathcal{V})\left[\frac{1}{2} - \frac{\mathcal{V}^2}{12}(2\tau\mathcal{V})^2\mathrm{e}^{2\varepsilon}\right] \geq \mathcal{V}^2\tau\left(1-\frac{2}{3}\varepsilon^2\mathrm{e}^{2\varepsilon}\right) \geq \frac{1}{2}\mathcal{V}^2\tau \\ &\geq \frac{1}{2}\left(\frac{t^d}{3^d}\right)^2 \frac{t}{3(2r)^\alpha} \geq \frac{1}{3^{1+2d}2^{1+\alpha}}\frac{t^{2d+1}}{r^\alpha}. \end{aligned} \tag{4.43}$$

This protocol shows that if the light cone of a Lieb-Robinson bound is $t \gtrsim r^\kappa$, then $\kappa \leq \alpha/(1+2d)$. □

Lastly, we note that it is trivial to remove the restriction $\dim(\mathcal{H}_i) = 2$ from the

assumptions of Theorem 3 by simply making $H(t)$ act trivially on all but 2 of the basis states in each $\mathcal{H}_i$.

4.4 The Frobenius Light Cone

We now turn to the Frobenius light cone. To motivate the development, let us consider the early time expansion of a time evolving operator $\mathcal{O}_i$, initially supported on lattice site i:

$$\mathcal{O}_i(t) = \sum_{n=0}^{\infty} \frac{(\mathcal{L}t)^n}{n!}\mathcal{O}_i = \mathcal{O}_i + \mathrm{i}t[H, \mathcal{O}_i] - \frac{t^2}{2}[H, [H, \mathcal{O}_i]] + \cdots \tag{4.44}$$

For illustrative purposes, we have temporarily assumed H is time-independent. Suppose further that H only contains nearest neighbor interactions. Then $[H, \mathcal{O}_i]$ can only contain operators of the form $\mathcal{O}_{i-1}\mathcal{O}\mathcal{O}_{i+1}$, and $[H, [H, \mathcal{O}_i]]$ can contain terms no more complicated than $\mathcal{O}_{i-2}\mathcal{O}_{i-1}\mathcal{O}\mathcal{O}_{i+1}\mathcal{O}_{i+2}$, and so on. It is natural to ask "how much" of the operator can be written as a sum of products of single-site operators restricted to some given subset of the lattice Λ. This question is naturally interpreted as follows: upon expanding $|\mathcal{O}_i(t))$ in terms of the basis vectors of Eq. (4.12):

$$|\mathcal{O}_i(t)) := \sum_{\{a_k\}} c_{\{a_k\}}(t) |\bigotimes_k T_k^{a_k}), \tag{4.45}$$

the coefficients $c_{\{a_k\}}(t)$ are analogous to the probability amplitudes of an ordinary quantum mechanical wave function. As we will see, the coefficients $c_{\{a_k\}}(t)$ must be sufficiently small if any a_k are non-identity, when the sites i and k are sufficiently far apart, at

any fixed time t: this is, intuitively, what we will call the Frobenius light cone.

For mathematical convenience in the discussion that follows, we restrict our analysis to finite lattices. It appears straightforward, if slightly tedious, to generalize to infinite lattices through an appropriate limiting procedure. More significantly, we will focus our discussion to one-dimensional lattices, as only in one dimension have we developed the machinery powerful enough to qualitatively improve upon the results of Section 4.3.

4.4.1 A Vector Space of Operators

We define a one dimensional lattice

$$\Lambda := \{i \in \mathbb{Z} : 0 \leq i \leq L\}. \tag{4.46}$$

For every site $i \in \Lambda$, we assume a finite dimensional local Hilbert space $\mathcal{H}_i$, obeying $\dim(\mathcal{H}_i) < \infty$. The global Hilbert space is

$$\mathcal{H} := \bigotimes_{i \in \Lambda} \mathcal{H}_i. \tag{4.47}$$

Let $\mathcal{B}$ denote the set of Hermitian operators acting on $\mathcal{H}$. We equip this space with the Frobenius inner product

$$(A|B) := \frac{\mathrm{tr}(AB)}{\dim(\mathcal{H})}, \tag{4.48}$$

upon which $\mathcal{B}$ becomes a real inner product space; we denote elements of this vector space $\mathcal{O} \in \mathcal{B}$ as $|\mathcal{O})$. When $A = B$, the inner product reduces to the squared Frobenius norm of A: $(A|A) = \|A\|_{\mathrm{F}}^2$. Note that for traceless operators A and B, this inner product

corresponds to the value of the thermal two-point connected correlation function at infinite temperature. Let $\{T_i^a\}$ denote the generators of $\mathrm{U}(\dim(\mathcal{H}_i))$, with $a = 0$ denoting the identity matrix. These generators form a complete basis for $\mathcal{B}$:

$$\mathcal{B} := \mathrm{span}\left\{\bigotimes_{i\in\Lambda} T_i^{a_i}\right\} := \mathrm{span}\left\{|a_0 \cdots a_L)\right\}. \tag{4.49}$$

We define the projectors

$$\mathbb{Q}_x|a_0 \cdots a_L) := \begin{cases} |a_0 \cdots a_L) & a_x \neq 0 \text{ and } a_y = 0 \text{ if } y > x \\ 0 & \text{otherwise} \end{cases}. \tag{4.50}$$

Hence $\mathbb{Q}_x$ selects the parts of an operator which all act on x, but on no site to the right of x. Observe that we have orthogonality and completeness:

$$\mathbb{Q}_i\mathbb{Q}_j = \delta_{ij}\mathbb{Q}_j, \qquad \sum_{i\in\Lambda} \mathbb{Q}_i = 1. \tag{4.51}$$

Time evolution is generated by a (generally time-dependent) Hamiltonian $H(t) : \mathbb{R} \to \mathcal{B}$. We assume that H is 2-local:

$$H(t) = \sum_{\{i,j\}\subset\Lambda} H_{ij}(t), \tag{4.52}$$

with power-law interactions of exponent α. By unitarity,

$$(\mathcal{O}|\mathcal{L}(t)|\mathcal{O}) = 0, \tag{4.53}$$

where $\mathcal{L}(t)$ was defined in Eq. (4.20); hence $\mathcal{L}(t)$ generates orthogonal transformations on $\mathcal{B}$ and leaves the length of all operators invariant.

4.4.2 The Operator Quantum Walk

Our goal is to understand the following scenario (Fig. 4.5): given an operator $|\mathcal{O})$ starting at the left-most site, i.e. obeying $\mathbb{Q}_0|\mathcal{O}) = |\mathcal{O})$, how long does it take before most of the operator $|\mathcal{O}(t))$ consists of operator strings that act on sites $\geq x$? More precisely, define

$$t_2^\delta(x) := \inf\left\{ t > 0 : \text{for any } \mathbb{Q}_0|\mathcal{O}_0) = |\mathcal{O}_0), \quad \delta < \frac{\sum_{y:x\leq y\leq L}(\mathcal{O}_0(t)|\mathbb{Q}_y|\mathcal{O}_0(t))}{(\mathcal{O}_0|\mathcal{O}_0)} \right\} \tag{4.54}$$

to be the shortest time for which a fraction δ of the operator $|\mathcal{O}(t))$ can be supported on sites $\geq x$. The assumption that the operator starts only on the left-most site is not restrictive—for an initial site $k \in \Lambda$, we can identify the lattice sites $k + m \sim k - m$ in order to "fold" the one dimensional lattice to put the initial point k at one boundary; such a change cannot modify Eq. (4.22), except to adjust the value of h by a factor < 4.

We note that

$$\begin{aligned} \sup_{\mathcal{O}_x\in\mathcal{B}_x} \frac{\mathrm{tr}\left([\mathcal{O}_0(t),\mathcal{O}_x]^\dagger[\mathcal{O}_0(t),\mathcal{O}_x]\right)}{\dim(\mathcal{H})(\mathcal{O}_0|\mathcal{O}_0)} &\leq 4\frac{(\mathcal{O}_0(t)|\mathbb{P}_x|\mathcal{O}_0(t))}{(\mathcal{O}_0|\mathcal{O}_0)} \\ &\leq 4\sum_{y:x\leq y\leq L}\frac{(\mathcal{O}_0(t)|\mathbb{Q}_y|\mathcal{O}_0(t))}{(\mathcal{O}_0|\mathcal{O}_0)}, \end{aligned} \tag{4.55}$$

where the left-most side corresponds to the out-of-time-order correlation function (OTOC) of an infinite-temperature state—a quantity known to herald the onset of many-body quan-

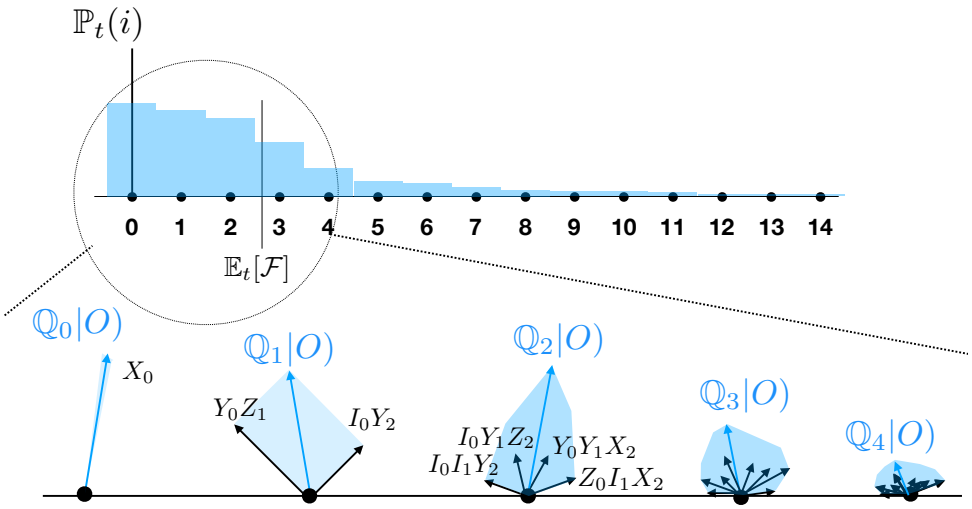

Figure 4.5: The 4^L dimensional space of operators can be decomposed into direct sum of L subspaces $\{\mathbb{Q}_i\}$ by the position of the right-most occupied site. By keeping track of only the "average value" of the right-most site (depicted above), keeping in mind that an exponential number of orthogonal operators (depicted below) are contained on most of the sites, we reduce the quantum walk of many-body operators from an exponentially large space to a one dimensional line.

tum chaos [90, 91]. From Eq. (4.55), it follows that a lower bound on $t_2^\delta(x)$ also bounds the evolution time of the OTOC and the growth of chaos.

The second main result of this chapter is the following theorem:

Theorem 4. *Given Hamiltonian evolution on $\mathcal{H}$ obeying Eq. (4.52) and Eq. (4.22), for any $x \in \Lambda$, $0 < \delta \in \mathbb{R}$ and $\frac{3}{2} < \alpha \in \mathbb{R}$, there exist constants $0 < K, K' < \infty$ such that*

$$t_2^\delta(x) \geq K \times \begin{cases} x & \alpha > \frac{5}{2} \\ x^{\alpha-3/2}(1 + K' \log x)^{-1} & \frac{3}{2} < \alpha \leq \frac{5}{2} \end{cases} . \tag{4.56}$$

Proof. We prove this theorem using the "operator quantum walk" formalism introduced in Ref. [94]. For simplicity, we will first prove the theorem when $\alpha > \frac{5}{2}$, and then

generalize to $\alpha \leq \frac{5}{2}$ afterwards. Consider the operator $\mathcal{F}$ acting on $\mathcal{B}$ defined by

$$\mathcal{F} := \sum_{j \in \Lambda} j \mathbb{Q}_j. \tag{4.57}$$

Our goal is to show that

$$\lim_{L \to \infty} \|[\mathcal{F}, \mathcal{L}(t)]\|_\infty \leq C < \infty. \tag{4.58}$$

The reason Eq. (4.58) is desirable is the following. Without loss of generality, we normalize $(\mathcal{O}|\mathcal{O}) = 1$. We then define a time-dependent probability distribution $\mathbb{P}_t$ on Λ as

$$\mathbb{P}_t(i \in \Lambda) := (\mathcal{O}(t)|\mathbb{Q}_i|\mathcal{O}(t)), \tag{4.59}$$

since by Eq. (4.51) the probability distribution is properly normalized: $\mathbb{P}_t(\Lambda) = 1$. We may then reinterpret $t_2^\delta(x)$ as the first time where the probability that $i \geq x$ on the measure $\mathbb{P}_t$ is sufficiently large:

$$t_2^\delta(x) = \inf \left\{ t > 0 : \delta < \mathbb{P}_t(i \geq x) \right\}. \tag{4.60}$$

We may then interpret $\mathcal{F}$ for $\alpha > \frac{5}{2}$ as a classical random variable that gives i with probability $\mathbb{P}_t(i)$. By Markov's inequality,

$$\mathbb{P}_t(i \geq x) \leq \frac{\mathbb{E}_t[\mathcal{F}]}{x}, \tag{4.61}$$

where $\mathbb{E}_t[\cdot]$ denotes expectation value on the measure $\mathbb{P}_t$. If Eq. (4.58) holds, then for any

operator $\mathcal{O}_0$ in the domain of $\mathbb{Q}_0$,

$$\begin{aligned}\mathbb{E}_t[\mathcal{F}] &= \int_0^t \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s}(\mathcal{O}_0(s)|\mathcal{F}|\mathcal{O}_0(s)) = \int_0^t \mathrm{d}s(\mathcal{O}_0(s)|[\mathcal{F},\mathcal{L}(s)]|\mathcal{O}_0(s)) \\ &\leq \int_0^t \mathrm{d}s\,|(\mathcal{O}_0(s)|[\mathcal{F},\mathcal{L}(s)]|\mathcal{O}_0(s))| \leq Ct. \end{aligned} \tag{4.62}$$

Combining Eq. (4.61) and Eq. (4.62), we see that Eq. (4.56) holds with

$$K = \frac{\delta}{C}. \tag{4.63}$$

Hence, it remains to prove Eq. (4.58). To do so, it will be useful to define

$$\widetilde{\Lambda} := \Lambda - \{0\}, \tag{4.64}$$

and a more refined set of complete, orthogonal projectors: for $S \subseteq \widetilde{\Lambda}$,

$$\mathbb{R}_S|a_0 \cdots a_L) := \begin{cases} |a_0 \cdots a_L) & i > 0 \text{ and } a_i \neq 0 \text{ if and only if } i \in S \\ 0 & \text{otherwise} \end{cases}, \tag{4.65}$$

which projects onto the operators whose support is exactly the subset S. We also define

$$\mathcal{F}_S := \max_{i \in S} i \tag{4.66}$$

to be the right-most occupied site. Observe that $\mathcal{F}_S \mathbb{R}_S = \mathbb{R}_S \mathcal{F} \mathbb{R}_S$. Since

$$\sum_{S \in \mathbb{Z}_2^{\tilde{\Lambda}}} \mathbb{R}_S = \mathbb{I}, \tag{4.67}$$

we may write, for any $\mathcal{O} \in \mathcal{B}$,

$$\begin{aligned}(\mathcal{O}|[\mathcal{F}, \mathcal{L}]|\mathcal{O}) &= \sum_{S,Q \in \mathbb{Z}_2^{\tilde{\Lambda}}} (\mathcal{O}|\mathbb{R}_S[\mathcal{F}, \mathcal{L}]\mathbb{R}_Q|\mathcal{O}) \\ &\leq \sum_{S,Q \in \mathbb{Z}_2^{\tilde{\Lambda}}} \sqrt{(\mathcal{O}|\mathbb{R}_S|\mathcal{O})(\mathcal{O}|\mathbb{R}_Q|\mathcal{O})} \sup_{\mathcal{O},\mathcal{O}' \in \mathcal{B}} \left| \frac{(\mathcal{F}_S - \mathcal{F}_Q)(\mathcal{O}|\mathbb{R}_S \mathcal{L} \mathbb{R}_Q|\mathcal{O}')}{\sqrt{(\mathcal{O}|\mathcal{O})(\mathcal{O}'|\mathcal{O}')}} \right|. \end{aligned} \tag{4.68}$$

Next, we observe that the 2-locality of the Hamiltonian implies that $\mathbb{R}_S \mathcal{L} \mathbb{R}_Q \neq 0$ if and only if there exists a site $i \in \Lambda$ such that $S \cup \{i\} = Q$ or $Q \cup \{i\} = S$.

Suppose that $Q \cup \{i\} = S$, that $\mathcal{F}_Q = j$ and that $i > 0$. Then if $i < j$, $\mathcal{F}_S = \mathcal{F}_Q = j$; the right-most occupied site in S and Q has not changed, and hence the supremum in Eq. (4.68) vanishes. Therefore, the supremum is only non-trivial when $i > j$. By submultiplicativity of the operator norm, there exists $0 < A < \infty$ such that

$$\begin{aligned}\sup_{\mathcal{O},\mathcal{O}' \in \mathcal{B}} \left| \frac{(\mathcal{F}_S - \mathcal{F}_Q)(\mathcal{O}|\mathbb{R}_S \mathcal{L} \mathbb{R}_Q|\mathcal{O}')}{\sqrt{(\mathcal{O}|\mathcal{O})(\mathcal{O}'|\mathcal{O}')}} \right| &\leq 2|i-j| \left\| \sum_{k \in Q} H_{ki} \right\|_\infty \\ &\leq 2|i-j| \sum_{k \in Q} \frac{h}{|i-k|^\alpha} \leq \frac{A}{|i - \mathcal{F}_Q|^{\alpha-2}}, \end{aligned} \tag{4.69}$$

where A is a constant and, in the last step, we overestimated the sum by assuming all sites $\leq j$ are included in the set Q. A similar argument holds when $S \cup \{i\} = Q$.

It is now useful to interpret Eq. (4.68) as an auxiliary linear algebra problem. Let

us define $\varphi \in \mathbb{R}^{\mathbb{Z}_2^{\tilde{\Lambda}}}$ as

$$\varphi_S := \sqrt{(\mathcal{O}|\mathbb{R}_S|\mathcal{O})}, \tag{4.70}$$

and $M \in \mathbb{R}^{\mathbb{Z}_2^{\tilde{\Lambda}}} \times \mathbb{R}^{\mathbb{Z}_2^{\tilde{\Lambda}}}$ as

$$M_{SQ} = M_{QS} := \begin{cases} A|\mathcal{F}_S - \mathcal{F}_Q|^{2-\alpha} & \mathcal{F}_S \neq \mathcal{F}_Q \text{ and } S = Q \cup \{m\} \text{ or } Q = S \cup \{m\} \\ 0 & \text{otherwise} \end{cases}. \tag{4.71}$$

Since

$$(\mathcal{O}|[\mathcal{F}, \mathcal{L}]|\mathcal{O}) \leq \sup_{\varphi : \|\varphi\|_{\mathrm{F}}=1} \sum_{S,Q} \varphi_S M_{SQ} \varphi_Q = \|M\|_\infty, \tag{4.72}$$

it suffices to show that $\|M\|_\infty < \infty$.

To bound the maximal eigenvalue of M, we use the min-max Collatz-Weiland Theorem [95]. To do that, we must first establish that M is an irreducible matrix (non-negativity of the entries is guaranteed by Eq. (4.71). To show irreducibility, we observe that

$$\left(M^{|S|}\right)_{\emptyset S} \neq 0; \tag{4.73}$$

the sequence of subsets which satisfies this identity corresponds to sequentially adding the elements of S from smallest to largest. We conclude that (by non-negativity of all M^n) there exists an $n \in \mathbb{Z}^+$ such that $(M^n)_{SQ} > 0$ for all sets S and Q.

We are now ready to apply the min-max Collatz-Weiland Theorem:

$$\|M\|_\infty = \inf_{\varphi \in \mathbb{R}^{\mathbb{Z}_2^{\tilde{\Lambda}}} : \varphi_S > 0} \max_S \frac{1}{\varphi_S} \sum_{Q \in \mathbb{Z}_2^{\tilde{\Lambda}}} M_{SQ} \varphi_Q. \tag{4.74}$$

Clearly an upper bound to the maximal eigenvalue comes from choosing any trial vector φ that we desire. We make the following choice: writing

$$S = \{n_1, \ldots, n_\ell\}, \text{ with } n_i < n_{i+1}, \tag{4.75}$$

we take $\varphi_\emptyset = 1$, and then define $n_0 = 0$ and

$$\varphi_S := \prod_{i=1}^{|S|} (n_i - n_{i-1})^{-\beta}, \tag{4.76}$$

where β is a tunable parameter we will shortly fix. Now we evaluate the right hand side of Eq. (4.74), defining $j = \mathcal{F}_S$:

$$\frac{1}{\varphi_S} \sum_{Q \in \mathbb{Z}_2^\Lambda} M_{SQ} \varphi_Q = M_{S,S-\{j\}} \frac{\varphi_{S-\{j\}}}{\varphi_S} + \sum_{k \in \Lambda : k > j} M_{S,S\cup\{k\}} \frac{\varphi_{S\cup\{k\}}}{\varphi_S} \tag{4.77}$$

Using Eq. (4.71), and assuming that $j_* = \mathcal{F}_{S-\{j\}}$,

$$M_{S,S-\{j\}} \frac{\varphi_{S-\{j\}}}{\varphi_S} \leq A(j - j_*)^{\beta+2-\alpha}. \tag{4.78}$$

We hence take

$$\beta = \alpha - 2 \tag{4.79}$$

to ensure that this first term is finite. Evaluating the second term of Eq. (4.77),

$$\sum_{k \in \Lambda : k > j} M_{S,S\cup\{k\}} \frac{\varphi_{S\cup\{k\}}}{\varphi_S} \leq A \sum_{k=j+1}^{\infty} (k - j)^{2-\alpha-\beta} \leq A_*, \tag{4.80}$$

where

$$A_* := A\frac{2^{\alpha+\beta-3}}{\alpha+\beta-3} < \infty, \tag{4.81}$$

so long as $\alpha > \frac{5}{2}$. We conclude that $C \leq A + A_* < \infty$, proving the theorem when $\alpha > \frac{5}{2}$.

We now return to the case $\frac{3}{2} < \alpha \leq \frac{5}{2}$. The proof is essentially identical with a few minor changes. Firstly, we set $\mathcal{F}_{\{0\}} = 0$, and for non-empty sets we define

$$\mathcal{F}_S := \max_{j \in S} \frac{j^\gamma}{1 + K' \log j}, \tag{4.82}$$

for a parameter $\gamma \in (0, 1)$ that we will fix shortly. We choose the parameter K' such that $\mathcal{F}_i$ is a convex function on $\mathbb{Z}^+$: $|\mathcal{F}_i - \mathcal{F}_j| \leq \mathcal{F}_{|i-j|}$. Such a K' can be shown to exist by extending $\mathcal{F}$ to act on $[1, \infty)$, after which we use elementary calculus to demand that

$$0 < \frac{\mathrm{d}\mathcal{F}(x)}{\mathrm{d}x} = \frac{1}{x^{1-\gamma}(1 + K' \log x)} \left(\gamma - \frac{K'}{1 + K' \log x} \right), \tag{4.83}$$

along with

$$\begin{aligned} 0 > \frac{\mathrm{d}^2\mathcal{F}(x)}{\mathrm{d}x^2} = -\frac{1}{x^{2-\gamma}(1 + K' \log x)} \Bigg[& \left(1 - \gamma + \frac{K'}{1 + K' \log x} \right) \\ & \times \left(\gamma - \frac{K'}{1 + K' \log x} \right) - \left(\frac{K'}{1 + K' \log x} \right)^2 \Bigg]. \end{aligned} \tag{4.84}$$

Eq. (4.83) and Eq. (4.84) are both satisfied by the choice

$$K' = \frac{\gamma}{4}. \tag{4.85}$$

We then find that convexity of $\mathcal{F}_i$ leads to the replacement of Eq. (4.71) with

$$M_{SQ} = M_{QS} := \begin{cases} A|\mathcal{F}_S - \mathcal{F}_Q|^{\gamma+1-\alpha}(1 + K' \log |\mathcal{F}_S - \mathcal{F}_Q|)^{-1} & \mathcal{F}_S \neq \mathcal{F}_Q \\ 0 & \text{otherwise} \end{cases}. \tag{4.86}$$

Lastly, we replace Eq. (4.76) with

$$\varphi_S := \prod_{i=1}^{|S|} \frac{(n_i - n_{i-1})^{\gamma+1-\alpha}}{1 + K' \log(n_i - n_{i-1})}. \tag{4.87}$$

These choices guarantee that

$$M_{S,S-\{j\}} \frac{\varphi_{S-\{j\}}}{\varphi_S} = A, \tag{4.88}$$

as in the prior setting. Then we find that

$$\sum_{k \in \Lambda : k > j} M_{S,S\cup\{k\}} \frac{\varphi_{S\cup\{k\}}}{\varphi_S} \leq A \sum_{k=j+1}^{\infty} \frac{1}{(k-j)^{2(\alpha-1-\gamma)}(1 + K' \log(k-j))^2}. \tag{4.89}$$

Upon choosing $\gamma = \alpha - \frac{3}{2}$, we obtain that the sum above is finite. Note that the logarithmic factors were required to obtain finiteness of Eq. (4.89). Hence we obtain $\|M\|_\infty < \infty$. Lastly, we mimic the arguments of Eq. (4.62) to complete the proof. □

We conjecture that in $d > 1$, the Frobenius light cone is always linear if and only if

$$\alpha > \frac{3d}{2} + 1. \tag{4.90}$$

We expect that for q-local Hamiltonians with $q > 2$, Eq. (4.90) holds only when a slightly

stricter requirement than Eq. (4.22) is obeyed: for example, if $\|H_{\{n_1,\ldots,n_q\}}\| \lesssim \prod_i |n_i - n_{i+1}|^{-\alpha}$ in one dimension.

The Frobenius light cone of Theorem 4 is tight up to subalgebraic corrections, when applied to arbitrarily large operators. This can be seen by considering a large operator of the form

$$\mathcal{O}_0 = \prod_{i=0}^{L/3} X_i^+ + \text{h.c.} \tag{4.91}$$

supported on the left-most $L/3$ sites of the lattice, where $X_i^+ = X_i + \mathrm{i}Y_i$. We would like to spread this operator to the right-most $L/3$ sites of the lattice, which are at least a distance $L/3$ away from the initial support. If the Hamiltonian is

$$H = \sum_{0\leq j\leq \frac{L}{3}} \sum_{\frac{2L}{3}<k\leq L} \frac{Z_j Z_k}{L^\alpha}, \tag{4.92}$$

it is straightforward to show that the fraction of $\mathcal{O}_0(t)$ supported beyond $2L/3$ is (up to the first order in t)

$$\sum_{k>L/2} \mathbb{Q}_k \mathcal{O}_0(t) = \frac{t}{3L^{\alpha-1}} \mathcal{O}_0 \sum_{\frac{2L}{3}<k\leq L} Z_k. \tag{4.93}$$

The Frobenius norm of this fraction is

$$\sqrt{\frac{L}{3}\left(\frac{t}{3L^{\alpha-1}}\right)^2} \propto \frac{t}{L^{\alpha-3/2}}. \tag{4.94}$$

Therefore, our bound in Theorem 4 is tight up to $O(1)$ factors.

4.5 The Free Light Cone

In this section, we discuss bounds on the quantum dynamics of non-interacting many-body systems.

4.5.1 Non-Interacting Hamiltonians

Consider a many-body quantum system defined on a d-dimensional lattice graph Λ; we assume the same properties of Λ as in Section 4.2. Suppose that the many-body Hamiltonian takes the form

$$H(t) = \sum_{i,j \in \Lambda} h_{ij}(t) c_i^\dagger c_j, \tag{4.95}$$

where $h(t) : \mathbb{R} \to \mathbb{C}^{\Lambda \times \Lambda}$ is a Hermitian matrix, and $c_i^\dagger$ and c_i represent either fermionic creation and annihilation operators:

$$\{c_j, c_i^\dagger\} := \delta_{ij}, \tag{4.96}$$

or bosonic creation and annihilation operators:

$$[c_j, c_i^\dagger] := \delta_{ij}. \tag{4.97}$$

The on site Hilbert space $\mathcal{H}_i$ obeys $\dim(\mathcal{H}_i) = 2$ in the fermionic case, and $\dim(\mathcal{H}_i) = \infty$ in the bosonic case. We note, however, that in isolated bosonic systems, $\mathcal{H}_i$ can often be truncated so that $\dim(\mathcal{H}_i)$ is at most the number of excitations on the lattice and is therefore finite.

As is well known, the evolution of all operators in such a non-interacting theory is controlled by the Green's function of the single particle problem on the Hilbert space $\mathbb{C}^{\Lambda}$. Time evolution on this space is generated by the Hamiltonian

$$H_{\mathrm{sp}}(t) := \sum_{i,j\in\Lambda} h_{ij}(t)|i\rangle\langle j|, \tag{4.98}$$

where $|i\rangle$ denotes the state that has exactly one excitation at site $i \in \Lambda$. The single particle time evolution matrix obeys the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}U_{\mathrm{sp}}(t) := -\mathrm{i}H_{\mathrm{sp}}(t)U_{\mathrm{sp}}(t), \tag{4.99}$$

together with the initial condition $U_{\mathrm{sp}}(0) = 1$. For example, in the fermionic model,

$$c_i(t) = \sum_{j\in\Lambda} U_{\mathrm{sp},ij}(t)c_j, \tag{4.100}$$

which follows from observing that

$$\frac{\mathrm{d}}{\mathrm{d}t}c_i := \mathrm{i}[H(t), c_i] = \sum_{j\in\Lambda} \mathrm{i}h_{ij}(t)[c_i^\dagger c_j, c_i] = -\mathrm{i}\sum_{j\in\Lambda} h_{ij}(t)c_j. \tag{4.101}$$

For simplicity in the discussion that follows, we drop the "sp" subscript on H and U.

4.5.2 Quantum Walks of A Free Particle

Consider a normalized wave function $|\psi(t)\rangle := U(t)|\psi\rangle \in \mathbb{C}^{\Lambda}$, along with its canonical probability distribution $\mathbb{P}_t$ on Λ:

$$\mathbb{P}_t(A) := \sum_{i \in A} \frac{|\langle i|\psi(t)\rangle|^2}{\langle \psi|\psi\rangle}. \tag{4.102}$$

Let us label an origin $0 \in \Lambda$, and assume that $|\psi(0)\rangle = |0\rangle$. We now use the quantum walk framework to prove our third main result, on the concentration of $\mathbb{P}_t$ on lattice sites close to the origin.

Theorem 5. *If $\alpha > d+1$, $\varepsilon > 0$, and $r \in \mathbb{Z}^+$, there exist constants $0 < K, u < \infty$ such that*

$$\sum_{y \in \Lambda : \mathcal{D}(y,x) \geq r} \mathbb{P}_t(y) \leq \frac{Kt}{(r-ut)^{\alpha-d-\varepsilon}}. \tag{4.103}$$

When $d < \alpha \leq d+1$, Eq. (4.103) holds with $u = 0$.

Proof. We first prove Eq. (4.103) when $\alpha > d+1$. Define the Hermitian operator

$$\langle x|\mathcal{F}(t)|y\rangle := \delta_{xy}\mathcal{F}(x,t), \tag{4.104a}$$

$$\mathcal{F}(x,t) := \min\left(0, \mathcal{D}(x,0) - ut\right). \tag{4.104b}$$

Our goal is to follow the proof of Theorem 4; first bounding the rate of change of an expectation value, and then employing Markov's inequality. The operator whose expectation value we will bound in the time evolved wave function is $\mathcal{F}^{\beta}$; ultimately we will see that $\beta = \alpha - d - \varepsilon$.

First, let us bound

$$\begin{aligned}\left|\mathcal{F}(x)^{\beta}-\mathcal{F}(y)^{\beta}\right| &\leq \beta \max(\mathcal{F}(x),\mathcal{F}(y))^{\beta-1}|\mathcal{F}(y)-\mathcal{F}(x)| \\ &\leq \beta \mathcal{D}(x,y)\left(\mathcal{F}(x)^{\beta-1}+\mathcal{F}(y)^{\beta-1}\right). \end{aligned} \tag{4.105}$$

Then,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\mathcal{F}^{\beta}|\psi(t)\rangle = -\mathrm{i}\langle\psi(t)|[\mathcal{F}^{\beta},H(t)]|\psi(t)\rangle - u\beta\langle\psi(t)|\mathcal{F}^{\beta-1}|\psi(t)\rangle. \tag{4.106}$$

Let us first bound the first term, using Eq. (4.105) and Eq. (4.106):

$$\begin{aligned}|\psi(t)|[\mathcal{F}^{\beta},H(t)]|\psi(t)\rangle| &\leq 2\sum_{\{x,y\}\subset\Lambda}|\langle x|[\mathcal{F}^{\beta},H(t)]|y\rangle|\,|\langle x|\psi\rangle\langle y|\psi\rangle| \\ &\leq \sum_{x\in\Lambda}\sum_{y\in\Lambda-\{x\}}(\mathbb{P}_t(x)+\mathbb{P}_t(y))|\langle x|[\mathcal{F}^{\beta},H(t)]|y\rangle| \\ &\leq \beta\sum_{x\in\Lambda}\sum_{y\in\Lambda-\{x\}}\mathbb{P}_t(x)\frac{2h}{\mathcal{D}(x,y)^{\alpha-1}}\left(\mathcal{F}(x)^{\beta-1}+\mathcal{F}(y)^{\beta-1}\right). \end{aligned} \tag{4.107}$$

In the last line, we have used the symmetry of the sum under exchanging x and y to remove $\mathbb{P}_t(y)$. Then we observe that

$$\mathcal{F}(y)^{\beta-1} \leq (\mathcal{F}(x)+\mathcal{D}(x,y))^{\beta-1} \leq 2^{\beta-1}\left(\mathcal{F}(x)^{\beta-1}+\mathcal{D}(x,y)^{\beta-1}\right). \tag{4.108}$$

Hence, so long as we choose

$$\beta = \alpha - d - \varepsilon, \tag{4.109}$$

we conclude that there exist constants $0 < K, A < \infty$ such that

$$
\begin{aligned}
&|\psi(t)|[\mathcal{F}^{\beta}, H(t)]|\psi(t)\rangle| \\
&\leq \left(2 + 2^{\beta}\right) \sum_{x \in \Lambda} \mathbb{P}_t(x) \sum_{y \in \Lambda - \{x\}} \left(\frac{h}{\mathcal{D}(x,y)^{\alpha-\beta}} + \frac{h}{\mathcal{D}(x,y)^{\alpha-1}} \mathcal{F}(x)^{\beta-1} \right) \\
&\leq \sum_{x \in \Lambda} \mathbb{P}_t(x) \left(K + A\mathcal{F}(x)^{\beta-1} \right) = K + A\langle \psi(t)|\mathcal{F}^{\beta-1}|\psi(t)\rangle, \qquad (4.110)
\end{aligned}
$$

where K, A are constants. Upon choosing

$$
u = \frac{A}{\beta}, \qquad (4.111)
$$

Eq. (4.106) implies that

$$
\langle \psi(t)|\mathcal{F}^{\beta}|\psi(t)\rangle \leq Kt. \qquad (4.112)
$$

Using Markov's inequality, and assuming $r > ut$,

$$
\sum_{y \in \Lambda : \mathcal{D}(y,x) \geq r} \mathbb{P}_t(y) \leq \frac{\mathbb{E}_t[\mathcal{F}^{\beta}]}{(r - ut)^{\beta}} \leq \frac{Kt}{(r - ut)^{\beta}}. \qquad (4.113)
$$

Combining Eq. (4.109) and Eq. (4.113) we obtain Eq. (4.103).

Secondly, we study the case $\alpha \leq d + 1$. Now we define

$$
\langle x|\mathcal{F}|y\rangle := \delta_{xy} \mathcal{D}(x,0)^{\beta}, \qquad (4.114)
$$

with β given by Eq. (4.109). Observe that $\beta < 1$. In this limit,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_t[\mathcal{F}] \leq \sum_{x\in\Lambda}\mathbb{P}_t(x)\sum_{y\in\Lambda-\{x\}}\frac{h}{\mathcal{D}(x,y)^\alpha}\left|\mathcal{D}(x,0)^\beta - \mathcal{D}(y,0)^\beta\right| \tag{4.115}$$

$$\leq \sum_{x\in\Lambda}\mathbb{P}_t(x)\sum_{y\in\Lambda-\{x\}}\frac{h}{\mathcal{D}(x,y)^{\alpha-\beta}}, \tag{4.116}$$

where in the last inequality we have used the convexity of $\mathcal{F}$ as a function of distance. For any $\varepsilon > 0$, the sum over y converges; hence there exists a $K < \infty$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_t[\mathcal{F}] \leq \sum_{x\in\Lambda}\mathbb{P}_t(x)\times K = K. \tag{4.117}$$

Another application of Markov's inequality implies Eq. (4.103). □

4.5.3 Free-Particle State Transfer

Our next goal is to prove the tightness of Theorem 5, up to subalgebraic corrections. This is achieved by the following theorem, which provides a rapid state-transfer protocol for a single particle.

Theorem 6. *For every $x \in \Lambda - \{0\}$ with $\mathcal{D}(x,0) > 2$, there exists a constant $0 < K < \infty$ and a Hermitian matrix $h(t) : \mathbb{R} \to \mathbb{C}^{\Lambda\times\Lambda}$ obeying Eq. (4.22), such that $\langle x|U(\tau)|0\rangle = 1$ at*

$$\tau := K \times \begin{cases} \mathcal{D}(x,0) & \alpha \geq d+1 \\ \mathcal{D}(x,0)^{\alpha-d} & d < \alpha < d+1 \\ \log\mathcal{D}(x,0) & \alpha = d \\ 1 & \alpha < d \end{cases}. \tag{4.118}$$

Proof. For $\alpha \geq d+1$, in order to transfer an excitation from 0 to x, we simply use a sequence of nearest-neighbor hoppings, which would take time proportional to the distance $\mathrm{dist}\,(()\,0, x)$. Specifically, let $(y_0 := 0, y_1, \ldots, y_{\ell-1}, y_\ell := x)$ be a sequence of length $1 + \mathcal{D}(x, 0)$ such that the edge (y_i, y_{i+1}) is an edge of nearest neighbors in Λ; here $\ell := \mathcal{D}(x, 0)$. We then apply

$$H(t) := \begin{cases} \mathrm{i}h|y_j\rangle\langle y_{j-1}| - \mathrm{i}h|y_{j-1}\rangle\langle y_j| & t \in [(j-1)\frac{\pi}{2h}, j\frac{\pi}{2h}) \\ 0 & \text{elsewhere} \end{cases}, \tag{4.119}$$

where h is defined in Eq. (4.22). It is straightforward to verify that the Hamiltonian takes $|y_{j-1}\rangle$ to $|y_j\rangle$ at the end of the interval $[(j-1)\frac{\pi}{2h}, j\frac{\pi}{2h}]$, for all $j = 1, \ldots, \ell$. As a result, we achieve perfect state transfer from site 0 to site x at $t = \frac{\pi}{2h}\mathrm{dist}\,(()\,0, x)$. Therefore, Theorem 6 holds for $\alpha > d+1$ with

$$K = \frac{\pi}{2h}. \tag{4.120}$$

For $\alpha < d+1$, we use a state-transfer scheme depicted in Fig. 4.3. The scheme, as depicted, assumes that the lattice is a simple cubic lattice; however, this protocol is easily applied to an arbitrary lattice graph, since we may always arrange the unit cells of the graph in the structure shown above. As the generalization to other lattices is obvious, we describe only the case of a simple cubic lattice below. We further assume the sites 0 and x are on the same axis of the lattice; more precisely, we assume that the path of shortest length connecting 0 and x is unique. If 0 and $x = (x_1, \ldots, x_d)$ do not satisfy this property, we use the protocol described below to to transfer the excitation from 0 to $(x_1, 0, \ldots, 0)$,

all the way to $(x_1, \ldots, x_d)$ in d separate steps, increasing the total transfer time by at most a factor of d compared to the protocol we describe below.

We define $q \in \mathbb{Z}^+$ as

$$q := \lfloor \log_2 \mathcal{D}(x, 0) \rfloor + 1. \tag{4.121}$$

Let $B_q \subset \mathbb{R}^d$ denote a cube of size $\mathcal{D}(x, 0)$ such that the sites 0 and x are at two different corners of B_q. We then recursively define a sequence of $q - 1$ cubes, namely $B_{q-1}^{(0)}, \ldots, B_1^{(0)}$, satisfying

$$\{0\} \subset B_1^{(0)} \subset B_2^{(0)} \subset \cdots \subset B_{q-1}^{(0)} \subset B_q^{(0)} := B_q, \tag{4.122}$$

and the size of $B_s^{(0)}$ is $2^s \frac{\text{dist}(()x,0)}{2^q}$ for all $s = 1, \ldots, q$. Note that our definition ensures the size of B_1 is in $[1, 2)$. Similarly, we define the cubes $B_1^{(x)}, \ldots, B_q^{(x)}$ that contains x:

$$\{x\} \subset B_1^{(x)} \subset B_2^{(x)} \subset \cdots \subset B_{q-1}^{(x)} \subset B_q^{(x)} := B_q. \tag{4.123}$$

Our strategy is to first expand the state $|\psi(0)\rangle = |0\rangle$ to a coherent uniform superposition on $B_1^{(0)}$, which is subsequently expanded to coherent uniform superpositions on larger and larger cubes $B_2^{(0)}, \ldots, B_q^{(0)}$. After that, we reverse the process and contract the uniform superposition on $B_q^{(0)} = B_q^{(x)}$ onto the cubes $B_{q-1}^{(x)}, \ldots, B_1^{(x)}$, and finally onto site $\{x\}$. We will argue that each expansion or contraction involving cubes of size ℓ takes time $\ell^{\alpha-d}$, where ℓ^α is the penalty due to the power-law constraint and ℓ^d is the enhancement coming from the volumes of the cubes. Summing over the values of ℓ results in a transfer time that scales as $\text{dist}\,(()\,0, x)^{\alpha-d}$ for $d < \alpha < d + 1$.

To calculate the time it takes for each expansion or contraction, we invoke the following Lemma:

Lemma 2. *Let A and B be two disjoint subsets of Λ, and $0 < C < \infty$. Then if*

$$|\psi(0)\rangle = \frac{1}{\sqrt{|A|}} \sum_{i \in A} |i\rangle, \tag{4.124}$$

there exists a free-particle Hamiltonian $H(t)$ defined in Eq. (4.95) with $|h_{ij}| \leq C$ for all $i, j \in \Lambda$ such that for any $\theta \in \mathbb{R}$,

$$|\psi(T)\rangle \propto \frac{\cos\theta}{\sqrt{|A|}} \sum_{i \in A} |i\rangle + \frac{\sin\theta}{\sqrt{|B|}} \sum_{i \in B} |i\rangle, \tag{4.125}$$

at time

$$T \leq \frac{\pi}{2C\sqrt{|B||A|}}. \tag{4.126}$$

Proof. We prove the lemma by construction. Consider the Hamiltonian

$$H(t) := \operatorname{sgn}(\tan\theta) \sum_{k \in A} \sum_{j \in B} \mathrm{i}C \left(|j\rangle\langle k| - |k\rangle\langle j| \right). \tag{4.127}$$

Without loss of generality, we take $\theta \in [0, \frac{\pi}{2}]$; the generalization to other θ is straightforward. By permutation symmetry, the wave function takes the form Eq. (4.125) with $\theta(t)$ a function of time. Pick any $j \in B$. We can explicitly evaluate

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\sqrt{|B|}}{\cos\theta} \frac{\mathrm{d}\langle j|\psi(t)\rangle}{\mathrm{d}t} = -\mathrm{i}\frac{\sqrt{|B|}}{\cos\theta} \langle j|H|\psi(t)\rangle = C\sqrt{|B||A|}. \tag{4.128}$$

Since the value of θ at which $|\psi(t)\rangle$ is given by Eq. (4.125) is in $[0, \frac{\pi}{2}]$, we conclude that Eq. (4.128) implies Eq. (4.126). □

By construction, the time τ of our perfect state transfer algorithm is given by

$$\tau = 2\sum_{s=1}^{q} T_{B,s}, \tag{4.129}$$

where $T_{B,s}$ is the time it takes to expand from $B_{s-1}^{(0)}$ to $B_s^{(0)}$, which is also the time it takes to contract $B_s^{(x)}$ into $B_{s-1}^{(x)}$. To evaluate these times, we use Lemma 2 with $C = 1/2^{s\alpha}$ to get

$$T_{B,s} \leq \pi \frac{2^{s\alpha}}{\sqrt{2^{(s-1)d}(2^{sd} - 2^{(s-1)d})}} = \frac{2^d \pi}{\sqrt{2^d - 1}} 2^{s(\alpha - d)}, \tag{4.130}$$

for all $s = 1, \ldots, q$. Here we have lower bounded the number of sites in $B_s^{(0)}$ by $2^{sd}/2$.

For $\alpha \neq d$, summing over s gives

$$\tau \leq \frac{2^{d+1}\pi}{\sqrt{2^d - 1}} \frac{2^{(q+1)(\alpha-d)} - 1}{2^{\alpha - d} - 1} \leq \begin{cases} \frac{2^{d+3}\pi}{\sqrt{2^d - 1}} \frac{1}{2^{\alpha-d} - 1} \text{dist}\left(() \, x, 0\right)^{\alpha - d} & (\alpha > d) \\ \frac{2^{d+1}\pi}{\sqrt{2^d - 1}} \frac{1}{1 - 2^{\alpha - d}} & (\alpha < d) \end{cases}. \tag{4.131}$$

On the other hand, at $\alpha = d$, we have

$$\tau \leq q \times \frac{2^d \pi}{\sqrt{2^d - 1}} \leq \frac{2^d \pi}{\sqrt{2^d - 1}} (1 + \log \text{dist}\left(() \, x, 0\right)). \tag{4.132}$$

Therefore, Theorem 6 follows. □

There are two important consequences of Theorem 6. Firstly, even a single quan-

tum mechanical degree of freedom can perform state transfer as asymptotically well as the previously best known protocol in an interacting many-body system [24] for $\alpha \geq d$. Secondly, Theorem 6 proves that any possible improvement to Theorem 5 must be sub-algebraic. Both the linear light cone and the superlinear polynomial light cones we have proved for free quantum systems with long-range interactions are now known to be optimal. Theorem 6 is also applicable to spin systems, since the spin degrees of freedom may be treated as hardcore bosons. Similarly, the protocol applies to Hamiltonians with on-site and particle number conserving interactions such as the Bose-Hubbard model: the interactions have no effect since at all times during the protocol there is at most a single particle in the system.

As noted in the introduction, this state-transfer protocol is naturally realized in experiments whenever there is a conserved quantity. For example, in a spin system with z-spin conservation, we can prepare the system in a highly polarized state with a single up-spin; the location of the up-spin represents the location of the single quantum degree of freedom, and our state-transfer protocol immediately applies. In trapped ion crystals, it is natural to use a large transverse magnetic field to help restrict to this subspace [96]. In addition, decoherence rates will be greatly reduced in the single-particle subspace, when compared to the GHZ states employed by Ref. [24].

A key feature of this state-transfer protocol is its remarkable robustness to error. Here we give a heuristic argument that this will be the case; a complete analysis will be provided elsewhere [97]. At step n of the protocol above, there are $\mathcal{N}_n = 2^{dn}$ sites in each domain which are all mutually coupled; the coherent state transfer process leads to an enhancement in the transfer rate by a factor of $\mathcal{N}_n$. Now suppose that there is uncorrelated

random error in the coefficients of $|j\rangle\langle k|$ in (4.127). Using random matrix theory [98], we estimate that these errors introduce lead to dephasing rates of order $\sqrt{\mathcal{N}_n}$. If $|x\rangle$ is the target site for the state-transfer protocol, we estimate the loss in fidelity $\mathcal{F} = |\langle\psi(\tau)|x\rangle|^2$ by summing up the error after each step:

$$1 - \mathcal{F} \sim \sum_{n=1}^{m} \frac{\varepsilon}{\sqrt{\mathcal{N}_n}} \sim \varepsilon \sum_{n=1}^{m} 2^{-dn/2} < \frac{\varepsilon}{2^{d/2} - 1}. \tag{4.133}$$

Here ε is related to the error in a single coupling in the state transfer process. Therefore, the quantum coherent hopping of this state-transfer protocol renders it highly immune to imperfections in tunable coupling constants which are inevitable in any near term quantum simulator. As $\varepsilon \to 0$, the fidelity $\mathcal{F} \to 1$.

4.6 Discussion

We have demonstrated a *hierarchy of linear light cones*—a sequence of metrics and protocols under which the emergent locality that arises in local quantum many-body systems breaks down at different exponents α of long-range interactions. The most general such light cone—the Lieb-Robinson light cone that bounds commutator norms—can become superlinear for any $\alpha < 2d + 1$. We conjecture that the Frobenius light cone that controls many-body chaos and state transfer can only be superlinear when $\alpha < 1 + \frac{3}{2}d$, and proved this result in $d = 1$ using the operator quantum walk formalism. Finally, in non-interacting systems, we proved both linear ($\alpha > d + 1$) and superlinear ($d < \alpha \leq d + 1$) light cones along with the optimality of these bounds. As such, we close a number of long-standing questions in the literature on the limits and capabilities of quantum dynam-

ics with long-range interactions.

Besides state transfer and many-body chaos, we have also demonstrated a wide range of applications of these (nearly) tight light cones. We proved that the growth of connected correlations obey the same light cone as that of the Lieb-Robinson bound. In the context of digital quantum simulation, we used the Lieb-Robinson bound to construct an approximation for the time-evolved version of a local observable, and thereby reduced cost of simulating the observable on quantum computers for all $\alpha > 2d + 1$. Similarly, using the free light cone, we constructed an efficient early-time classical boson sampler for all $\alpha > d$, exponentially improving the previous best estimate in some regime of α. Additionally, we bounded the time it takes to generate topologically ordered states using power-law interactions. Finally, we tightened the minimum correlation-decay rate in the ground state of a gapped power-law Hamiltonian.

The hierarchy of linear light cones revealed in this chapter has important implications both on the capabilities of quantum technologies exploiting long-range interactions, as well as on the nature of quantum information dynamics and thermalization in these systems. A complete understanding of quantum chaos and state transfer, at the very least, requires the construction of a new mathematical framework beyond the Lieb-Robinson bounds, perhaps along the lines of our operator quantum walk. It is an interesting open question whether and how the hierarchy of different notions of locality revealed in this manuscript reveals itself in aspects of quantum chaos besides OTOCs, perhaps including entanglement dynamics or eigenvalue statistics. Lastly, it also remains an important future challenge to obtain the Frobenius light cone in two or more dimensions, as well as to rigorously study the light cone that controls the decoherence of a quantum system subject

to long-range random noise, which was conjectured to be linear for $\alpha > d + \frac{1}{2}$ [99].

Chapter 5: Locality and Digital Quantum Simulation of Power-Law Interactions

Lieb-Robinson bounds limit the rate at which information can propagate in systems that obey the laws of non-relativistic quantum mechanics [12, 18, 18, 19, 27, 28, 29, 30, 31, 32, 33]. These bounds have found a plethora of applications [36, 100, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110], including recent results on entanglement area laws [111, 112, 113], the classical complexity of sampling bosons [16], and even a quantum algorithm for digital quantum simulation [76].

Lieb and Robinson's original proof applies only to short-range interactions, i.e., those that act over a finite range or decay at least exponentially in space. However, interactions in many physical systems, such as trapped ions [49, 50], Rydberg atoms [55], ultracold atoms and molecules [53, 59], nitrogen-vacancy centers [58], and superconducting circuits [114], can decay with distance r as a power law ($1/r^\alpha$) and, hence, lie outside the scope of the original Lieb-Robinson bound. Thus, understanding the fundamental limit on the speed of information propagation in these systems holds serious physical implications, including for the applications mentioned above. Despite many efforts in recent years [18, 19, 29, 30], a *tight* Lieb-Robinson bound for such long-range interactions remains elusive.

In this chapter, we derive a new Lieb-Robinson bound for systems with power-law decaying interactions in D dimensions. While our bound is not known to be tight, it has four main benefits compared to the best previous bound for such systems [19]: (i) It is tighter, resulting in the best effective light cone to date [Eq. (5.17)]. (ii) The bound applies at all times, and not just asymptotically in the large-time limit. (iii) The framework behind the proof is conceptually simpler, with an easy-to-understand interpretation based on physical intuition. (iv) Our approach is potentially applicable to studying a wider variety of quantities, including connected correlators [36, 115] and higher-order correlators (for instance, the out-of-time-ordered correlator [116, 117] and the full measurement statistics of boson sampling [16, 82]) as we discuss in Sec. 5.5.

In contrast to the previous long-range Lieb-Robinson bounds [18, 19, 29, 30], which all relied on the so-called Hastings-Koma series [18], our approach is based on a generalization of the framework Haah et al. [76] (HHKL) introduced as a building block for their quantum simulation algorithm. The essence of their framework is a technique for decomposing the time evolution of a system into evolutions of subsystems, with an error bounded by the Lieb-Robinson bound for short-range interactions [12]. We extend the HHKL framework to long-range interactions and to a more general choice of subsystems. Remarkably, these modifications enable us to derive a tighter Lieb-Robinson bound for long-range interactions than the one we use in the analysis of the decomposition [29].

Additionally, we return to the original motivation of Haah et al.'s framework: the digital simulation of lattice-based quantum systems. We generalize the HHKL algorithm to simulate systems with power-law decaying interactions. The algorithm scales better as a function of system size than previous algorithms when $\alpha > 3D$, and the speed-up

becomes more dramatic as α is increased.

The structure of the chapter is as follows. In Sec. 5.1, we state our main results and summarize the proof of the new Lieb-Robinson bound. In Sec. 5.2, we lay out the precise mathematical framework for the proof and generalize the technique for decomposing time-evolution unitaries [76] to power-law decaying interactions and to more general choices of subsystems. After that, we present two applications of the unitary decomposition in Sec. 5.3 and Sec. 5.4, which can be read independently of each other. Specifically, in Sec. 5.3, we use the unitary decomposition to derive the improved Lieb-Robinson bound for long-range interactions. In Sec. 5.4, we analyze the performance of the HHKL algorithm from Ref. [76] when applied to simulating long-range interacting systems. We conclude in Sec. 5.5 with an outlook for the future.

5.1 Overview of Results

In this section, we summarize our main results for the case of a one-dimensional lattice. Without loss of generality, we assume that the distance between neighboring sites is one. The unitary decomposition technique in Sec. 5.2 is generalized from a similar result for short-range interactions in Ref. [76]. We use it to approximate the evolution of a long-range interacting system ABC by three sequential evolutions of its subsystems AB, B, and BC (see Fig. 5.1). We assume that the interaction strength between any two sites in the system is bounded by $1/r^{\alpha}$, with r being the distance between the sites and α a nonnegative constant. This restriction on the Hamiltonian norm also sets the time unit for the evolution of the system.

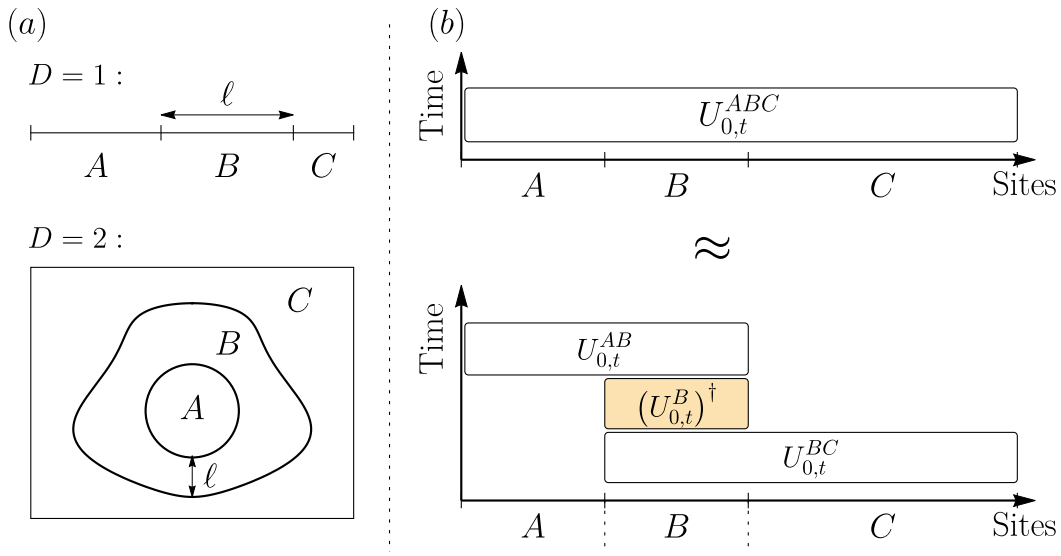

Figure 5.1: A demonstration of the unitary decomposition in Lemma 3. Panel (a): the three disjoint regions A, B, C in $D = 1$ and $D = 2$ dimensions with A convex and compact. Panel (b): Lemma 3 allows the evolution of the whole system to be approximated by a series of three evolutions of subsystems. The horizontal axis lists the sites in each of the three sets A, B, C (not necessarily according to their geometrical arrangement, particularly in higher dimensions). Each box is an evolution for time t of a Hamiltonian supported on the sites the box covers. These evolutions can be forward (white fill) or backward (orange fill, with dagger) in time.

There are two sources of error in the approximation: one due to the truncation of the Hamiltonian of the system ABC (we ignore the interactions that connect A and C), and the other due to the Hamiltonians of the subsystems AB, B, and BC not commuting with each other. For a fixed value of α, if the distance ℓ between the two regions A and C (see Fig. 5.1a) is large enough, namely $\ell \gg \alpha$, the two error sources have the same scaling with ℓ. To estimate the error, for example from the truncation, we sum over interactions connecting sites in A and C, and obtain a total error of $\mathcal{O}\left(1/\ell^{\alpha-2}\right)$ (in one dimension) for the approximation in the unitary decomposition (as shown in Appendix B.1.1).

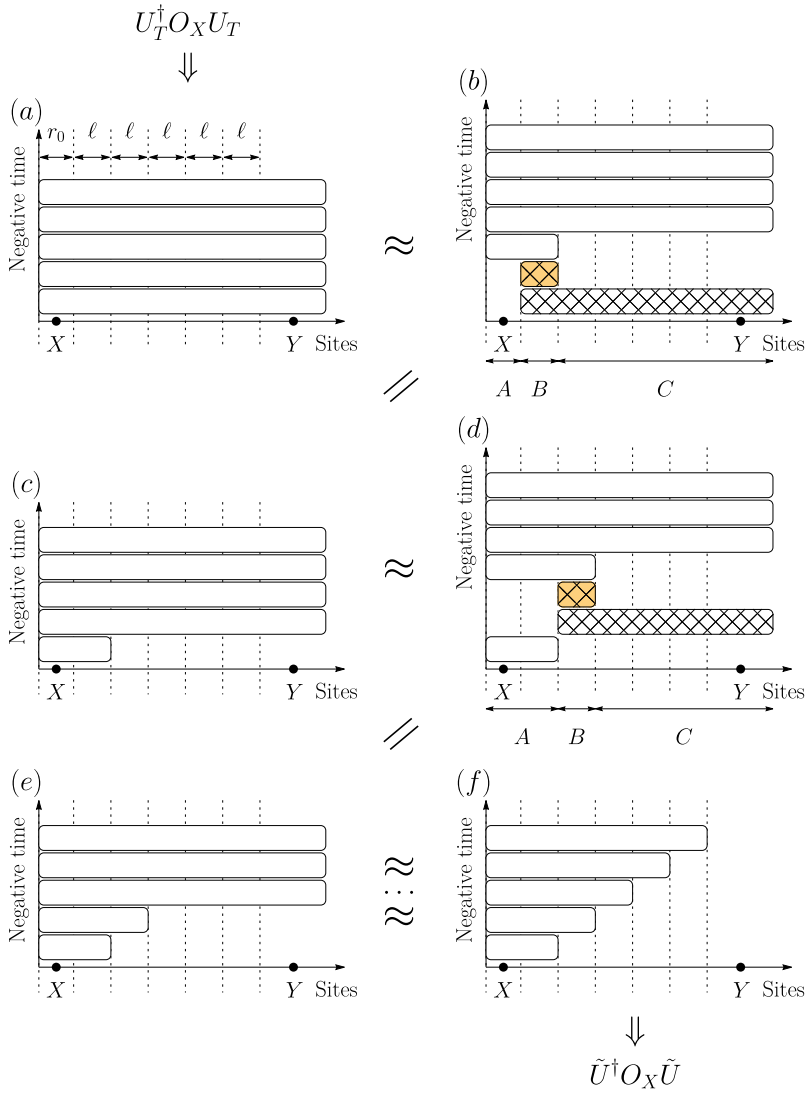

Figure 5.2: A step-by-step construction of the unitary $\tilde{U}$ such that $\tilde{U}^\dagger O_X \tilde{U} \approx U_T^\dagger O_X U_T$. Each box represents an evolution of the subsystem covered by the width of the box for a fixed time. The colors of the boxes follow the same convention as in Fig. 5.1. In panel *(a)*, the unitary U_T is written as a product of evolutions of the same system in $M = 5$ consecutive time slices. *(b)* The evolution in the last (bottom) time slice is decomposed using the method in Fig. 5.1, with the choice of subsystems A, B, C such that X is contained in A. The evolutions of the subsystems B and BC (hatched boxes) therefore commute with O_X and cancel out with their counterparts from $U_T^\dagger$, resulting in *(c)*. In panel *(d)*, we repeat the procedure for the second-from-bottom time slice, but note the different choice of A, B, C from panel *(b)*. This difference is necessary to ensure that the evolutions of B and BC commute with the evolution(s) from the previously decomposed time slice(s). We then commute them through O_X again and remove them from the construction of $\tilde{U}$ in panel *(e)*. Repeatedly applying the unitary decomposition for the other time slices, we obtain the unitary $\tilde{U}$ in panel *(f)*, which is supported on a smaller region than the original unitary U_T. With a proper choice of the size ℓ of B, we can make sure that Y lies outside this region, and, therefore, $\tilde{U}$ commutes with O_Y.

In Sec. 5.3, we use the unitary decomposition to prove a Lieb-Robinson bound for long-range interactions that is stronger than previous bounds, including the one we use in the proof of the unitary decomposition. The subject of such a bound is usually the norm of the commutator $\|[O_X(T), O_Y]\|$ between an operator $O_X(T) = U_T^\dagger O_X U_T$ evolved under a long-range Hamiltonian for time T and another operator O_Y supported on a set Y that is at least a distance R away from the support X of O_X. Here, we briefly explain the essence of the proof using a one-dimensional system with fixed α and large enough $R, T \gg \alpha$ as an example. The strategy is to use the aforementioned unitary decomposition to construct another unitary $\tilde{U}$ such that *(i)* $\tilde{U}^\dagger O_X \tilde{U}$ approximates $U_T^\dagger O_X U_T$ and *(ii)* $\tilde{U}^\dagger O_X \tilde{U}$ commutes with O_Y, so the commutator norm $\|[O_X(T), O_Y]\|$ will be approximately zero, up to the error of our approximation. For fixed α, we consider $M \propto T$ equal time slices and use the unitary decomposition to extract the relevant parts from the evolution U_T in each time slice. Each time we decompose a unitary, we choose the subsystems A, B, C so that only A overlaps with the supports of the unitaries from the previous time slices (see Fig. 5.2), and therefore the evolutions of B and BC can be commuted through O_X to cancel their counterparts from $U_T^\dagger$ (Fig. 5.2b and Fig. 5.2d):

$$\begin{aligned}
&\left(U^{ABC}\right)^\dagger O_X U^{ABC} \\
&\approx \left(U^{AB}\right)^\dagger U^B \left(U^{BC}\right)^\dagger O_X U^{BC} \left(U^B\right)^\dagger U^{AB} \\
&= \left(U^{AB}\right)^\dagger O_X U^{AB}.
\end{aligned} \tag{5.1}$$

The remaining evolutions that contribute to the construction of $\tilde{U}$ are supported entirely on a ball of radius $\sim M\ell$ around X, where ℓ is the size of B and is chosen to be the

same in all time slices. By choosing $\ell \sim R/M$ and $M\ell < R$ so that Y lies outside this ball, the commutator norm $\|[O_X(T), O_Y]\|$ is at most the number of time slices multiplied by $\mathcal{O}\left(1/\ell^{\alpha-2}\right)$, which is the decomposition error per time slice. Therefore, we obtain a Lieb-Robinson bound for long-range interactions in one dimension:

$$\|[O_X(T), O_Y]\| \leq c_{\mathrm{lr},\alpha} \frac{T}{\ell^{\alpha-2}} = c_{\mathrm{lr},\alpha} \frac{T^{\alpha-1}}{R^{\alpha-2}}, \tag{5.2}$$

where $c_{\mathrm{lr},\alpha}$ is a constant that may depend on α, but not on T, R. Setting the commutator norm to a small constant yields the causal region inside the effective light cone: $T \gtrsim R^{\frac{\alpha-2}{\alpha-1}}$. For comparison, the previous best Lieb-Robinson bound produces a light cone $T \gtrsim R^{\frac{\alpha-2}{\alpha}}$ [19]. Our bound is therefore tighter in the asymptotic limit of large R and large T, while its proof is substantially more intuitive than in Ref. [19]. A more careful analysis (Sec. 5.3) shows that our light cone also becomes linear in the limit $\alpha \to \infty$, where the power-law decaying interactions are effectively short-range. Moreover, our bound works for arbitrary time T, while the bound in Ref. [19] applies only in the long-time limit. We provide a more rigorous treatment as well as a bound for D-dimensional systems in Sec. 5.3.

Section 5.4 then then discusses the original motivation for the unitary decomposition—digital quantum simulation—in the case of long-range interactions that decay as a power law. For $\alpha > 2D$, our analysis shows that the HHKL algorithm [76] requires only $\mathcal{O}\left(Tn(Tn/\varepsilon)^{\frac{2D}{\alpha-D}} \log \frac{Tn}{\varepsilon}\right)$ two-qubit gates to simulate the evolution of a system of n sites arranged in a D-dimensional lattice for time T with an error at most ε. For large α, the gate count of the algorithm scales with n significantly better than other algorithms.

5.2 Mathematical Framework

In this section, we present the technique for approximating the time evolution of a system by evolutions of subsystems. We later use this technique to derive a stronger Lieb-Robinson bound (Sec. 5.3) and an improved quantum simulation algorithm (Sec. 5.4) for systems with long-range interactions.

We consider n sites arranged in a D-dimensional lattice $\Lambda \subset \mathbb{N}^D$ of size $L = \mathcal{O}\left(n^{1/D}\right)$ and $D \geq 1$. Recall that, without loss of generality, we assume the spacing between neighboring lattice sites is one. This assumption sets the unit for distances between sites in the lattice. We shall embed the lattice Λ into the real space $\mathbb{R}^D$. The intersection $X \cap \Lambda$ therefore contains every lattice site in a subset $X \subset \mathbb{R}^D$. The system evolves under a (possibly) time-dependent Hamiltonian $H_\Lambda(t) = \sum_{\vec{i},\vec{j}} h_{\vec{i},\vec{j}}(t)$, with $h_{\vec{i},\vec{j}}(t)$ being the interaction between two sites $\vec{i}, \vec{j} \in \Lambda$. Without ambiguity, we may suppress the time-dependence in the Hamiltonians. We say a system has power-law decaying interactions if $\|h_{\vec{i},\vec{j}}\| \leq \frac{1}{\|\vec{i}-\vec{j}\|^\alpha}$, where $\|\cdot\|$ denotes both the matrix and the vector 2-norms, for some nonnegative constant α and for all $\vec{i} \neq \vec{j}$. [Note that $h_{\vec{i},\vec{i}}$ may have arbitrarily large norm.] For readability, we denote by $H_X = \sum_{\vec{i},\vec{j} \in X} h_{\vec{i},\vec{j}}$ the terms of H_Λ that are supported entirely on a subset $X \cap \Lambda$, and by $U^X_{t_1,t_2} \equiv \mathcal{T} \exp\left(-i \int_{t_1}^{t_2} H_X dt\right)$ the evolution unitary under H_X from time t_1 to t_2, where $\mathcal{T}$ is the time-ordering operator. We also denote by $\mathrm{dist}\,(X, Y)$ the minimum distance between any two sites in X and Y, by $X^c = \mathbb{R}^D \setminus X$ the complement of X in real space, by ∂X the boundary of a compact subset X, by $\Phi(X)$ the area of ∂X, and by XY the union $X \cup Y$. In the following, we keep track of how errors scale with time, distance, and α, while treating the dimension D as a constant.

We now describe how to approximate the evolution of the system to arbitrary precision by a series of evolutions of subsystems using a technique we generalize from Ref. [76].

Lemma 3. *Let $A, B, C \subset \mathbb{R}^D$ be three distinct regions with non-empty interiors such that $A \cup B \cup C = \mathbb{R}^D$. Let A be both compact (closed and bounded) and convex. We have*

$$\left\| U_{0,t}^{ABC} - U_{0,t}^{AB} \left(U_{0,t}^{B}\right)^{\dagger} U_{0,t}^{BC} \right\| \leq c_0 (e^{vt} - 1) \Phi(A) \xi_\alpha(\ell),$$

with

$$\xi_\alpha(\ell) = \left(\frac{16}{1-\gamma}\right)^{\alpha} \frac{1}{\ell^{\alpha - D - 1}} + e^{-\gamma \ell}, \tag{5.3}$$

for all $\alpha > D + 1$. Here, $v, c_0 \in \mathbb{R}^+$ are positive constants, γ is a constant that can be chosen arbitrarily in the range $(0, 1)$, and $\ell = dist\,(A, C)$ is the distance between sets A and C.

We emphasize that this lemma applies to arbitrary sets A that are both convex and compact. The sets we focus on include D-balls and hyperrectangles in $\mathbb{R}^D$. The former geometry is relevant in the proof of our new Lieb-Robinson bound, the latter in the analysis of the HHKL algorithm for long-range interactions.

Lemma 3 allows us to approximate the evolution of a long-range interacting system ABC by that of subsystems AB, B, BC (Fig. 5.1). The features of the function $\xi_\alpha(\ell)$ are better understood by considering two limiting cases of physical interest. First, when α is finite and ℓ (the distance between A and C) is large compared to α, the function $\xi_\alpha(\ell)$

behaves like

$$\mathcal{O}\left(\frac{1}{\ell^{\alpha-D-1}}\right), \tag{5.4}$$

which decays only polynomially with ℓ. In the second limit, as $\alpha \to \infty$ for a large but finite ℓ, we recover from $\xi_\alpha(\ell)$ the exponentially decaying error bound $e^{-\gamma\ell}$—a trademark of finite-range interactions [12, 76].

The proof of Lemma 3, while more general, bears close resemblance to the corresponding analysis for short-range interactions in Ref. [76]. However, there are two key differences. First, in order to make the approximation in Lemma 3, some interactions between sites separated by a distance greater than ℓ are truncated from the Hamiltonian. While such terms vanish in a system with short-range interactions, here they contribute $\mathcal{O}\left(\Phi(A)/\ell^{\alpha-D-1}\right)$ to the error of the approximation. In addition, instead of the original Lieb-Robinson bound [12] which applies only to systems with short-range interactions, we use Gong et al.'s generalization of the bound for long-range interactions [29]. The result is an approximation error that decays with ℓ polynomially as $\mathcal{O}\left(\Phi(A)/\ell^{\alpha-D-1}\right)$, *in addition* to the exponentially decaying error that exists already for short-range interactions. Nevertheless, the error can always be made arbitrarily small by choosing ℓ to be large enough.

In Sec. 5.2.1 below, we present the proof of Lemma 3. After that, we demonstrate the significance of Lemma 3 with two applications: a stronger Lieb-Robinson bound for long-range interacting systems (Sec. 5.3) and an improved error bound for simulating these systems (Sec. 5.4). Both sections are self-contained, and readers may elect to focus

on either of them.

5.2.1 Error bound on the unitary decomposition

Here, we will outline the proof of Lemma 3. Similar to Ref. [76], we begin with an identity:

$$U_{0,t}^{ABC} = U_{0,t}^{AB} U_{0,t}^{C} \underbrace{\left(U_{0,t}^{C}\right)^{\dagger} \left(U_{0,t}^{AB}\right)^{\dagger} U_{0,t}^{ABC}}_{=W_t}. \tag{5.5}$$

Our aim is to approximate W_t by $\left(U_{0,t}^{C}\right)^{\dagger} \left(U_{0,t}^{B}\right)^{\dagger} U_{0,t}^{BC}$, from which Lemma 3 will follow. For that, we look at the generator of W_t [76], i.e., a Hamiltonian $\mathcal{G}_t$ such that

$$\frac{dW_t}{dt} = -i\mathcal{G}_t W_t, \tag{5.6}$$

for all time. Exact differentiation of W_t yields [89, 118]

$$\mathcal{G}_t = \left(U_{0,t}^{C}\right)^{\dagger} \left(U_{0,t}^{AB}\right)^{\dagger} (\underbrace{H_{ABC} - H_{AB} - H_C}_{=H_{A:C}+H_{B:C}}) U_{0,t}^{AB} U_{0,t}^{C} \tag{5.7}$$

$$= \left(U_{0,t}^{C}\right)^{\dagger} \left(U_{0,t}^{AB}\right)^{\dagger} H_{B:C} U_{0,t}^{AB} U_{0,t}^{C} + \delta_{\text{trunc}} \tag{5.8}$$

$$= \left(U_{0,t}^{C}\right)^{\dagger} \left(U_{0,t}^{B}\right)^{\dagger} H_{B:C} U_{0,t}^{B} U_{0,t}^{C} + \delta_{\text{overlap}} + \delta_{\text{trunc}}, \tag{5.9}$$

where $H_{X:Y} = \sum_{i\in X, j\in Y} h_{ij}(t)$ denotes the sum of terms supported *across* disjoint sets X and Y, and $\delta_{\text{trunc}}, \delta_{\text{overlap}}$ are error terms we now define and evaluate. Note that the first term in Eq. (5.9) is the generator of $\left(U_{0,t}^{C}\right)^{\dagger} \left(U_{0,t}^{B}\right)^{\dagger} U_{0,t}^{BC}$—the unitary with which we aim to approximate W_t.

In contrast to the approximation for short-range interacting systems in Ref. [76], there are two sources of error in Eq. (5.9). The first error term δ_{trunc} arises after we discard $H_{A:C}$ from Eq. (5.7). For the short-range interactions in Ref. [76], this error vanishes when the distance ℓ between A and C is larger than the interaction range. However, in our case, there is a nontrivial truncation error associated with ignoring long-range interactions between A and C:

$$\|\delta_{\text{trunc}}\| = \|H_{A:C}\| = c_{\text{tr}} 2^{\alpha} \frac{\Phi(A)}{\ell^{\alpha - D - 1}} \tag{5.10}$$

for $\alpha > D + 1$, where c_{tr} is a constant [Eq. (B.5)], $\ell = \text{dist}\,(A, C)$ is the distance between A and C. The factor of $1/l^{\alpha}$ in the bound comes from the requirement that the two-body interactions decay as a power law $1/r^{\alpha}$, while the term ℓ^{D} is due to the sum over all sites in the D-dimensional set C. Another factor of $\ell\Phi(A)$ arises after summing over the volume of A, which we assume to be a compact and convex set. The detailed evaluation of the norm is presented in Appendix B.1.1.

The other error, which we define to be δ_{overlap}, is the result of the approximation used between Eqs. (5.8) and (5.9). In the former equation, the operator evolves under $H_{AB} + H_C$, whereas in the latter, it evolves under the reduced Hamiltonian $H_B + H_C$, thus incurring the error:

$$\|\delta_{\text{overlap}}\| = \left\| \left(U_{0,t}^{AB}\right)^{\dagger} H_{B:C} U_{0,t}^{AB} - \left(U_{0,t}^{B}\right)^{\dagger} H_{B:C} U_{0,t}^{B} \right\|. \tag{5.11}$$

To understand why $\|\delta_{\text{overlap}}\|$ is small, recall that $H_{B:C}$ is the sum of terms $h_{\vec{b},\vec{c}}$ that are

supported on two sites $\vec{b} \in B$ and $\vec{c} \in C$. Since the strengths of such terms decay as $1/r^\alpha$ (with r the distance between the sites $\vec{b}$ and $\vec{c}$), the main contribution to $H_{B:C}$—and thus to δ_{overlap}—comes from the terms where $\vec{b}$ and $\vec{c}$ are spatially close to each other. But since the sets A, C are separated by a large distance ℓ, if the site $\vec{b}$ is close to C, then it must be far from A. Thus, the evolution of $h_{\vec{b},\vec{c}}$ for a short time under H_{AB} can be well-approximated by evolution under H_B alone. In Appendix B.1.2, we make this intuition rigorous using Gong et al. [29]'s generalization of the Lieb-Robinson bound to systems with long-range interactions.

In the end, we obtain the following bound on δ_{overlap}:

$$\|\delta_{\text{overlap}}\| \leq c_{\text{ov}}(e^{vt} - 1)\Phi(A)\left[\frac{\left(\frac{16}{1-\gamma}\right)^\alpha}{\ell^{\alpha-D-1}} + \frac{1}{e^{\gamma\ell}}\right], \tag{5.12}$$

where c_{ov} is a constant [Eq. (B.24)] and $\gamma \in (0, 1)$ is a free parameter. The bound has contributions from two competing terms: one that decays polynomially with ℓ and another that decays exponentially. The polynomially decaying term is dominant for fixed α and large ℓ, whereas the exponentially decaying term prevails as $\alpha \to \infty$ for fixed ℓ. The errors δ_{trunc} and δ_{overlap} in approximating the generator $\mathcal{G}_W$ combine to give an overall error in approximating W_t with $\left(U^C_{0,t}\right)^\dagger \left(U^B_{0,t}\right)^\dagger U^{BC}_{0,t}$ (see Appendix B.2). From this, we obtain the error bound in Lemma 3, with $c_0 = \max\{c_{\text{tr}}, c_{\text{ov}}\}/v$.

Before discussing applications of Lemma 3, we pause here to note that the Lieb-Robinson bound in Gong et al. [29] used in the above analysis is not the tightest-known bound for long-range interactions [19]. Our use of this bound, however, does not lead to a suboptimal error bound in Lemma 3. For finite α, the error bound is dominated by the

polynomially decaying term $1/\ell^{\alpha-D-1}$, which arises from the truncation error δ_{trunc} rather than δ_{overlap}. Therefore, this error term would not benefit from a tighter Lieb-Robinson bound. In the limit $\alpha \to \infty$, on the other hand, we shall see later that the lemma already reproduces the short-range Lieb-Robinson bound, which is optimal up to a constant factor. Thus, we expect that using stronger Lieb-Robinson bounds would produce no significant improvement for the error bound in Lemma 3.

5.3 A Stronger Lieb-Robinson Bound from Quantum Simulation Algorithms

In this section, we will use Lemma 3 to derive a stronger Lieb-Robinson bound for long-range interactions. The first generalization of the Lieb-Robinson bound to power-law decaying interactions was given by Hastings and Koma [18]. However, their bound diverges in the limit $\alpha \to \infty$, where the power-law decaying interactions are effectively short-range. Later, Gong et al. [29] derived a different bound that, in this limit, does indeed converge to the Lieb-Robinson bound for short-range interactions. While we used this bound in Sec. 5.2 to prove Lemma 3, we will also show that by *using* this lemma, we can in turn derive a Lieb-Robinson bound for long-range interactions that is stronger than the one in Gong et al. In fact, our bound produces a tighter effective light cone than even the strongest Lieb-Robinson bound for long-range interactions known previously [19].

Recall that the subject of a Lieb-Robinson bound is the commutator norm

$$\mathcal{C}(T,R) \equiv \left\| \left[\left(U_{0,T}^{\Lambda}\right)^{\dagger} O_X U_{0,T}^{\Lambda}, O_Y \right] \right\|, \tag{5.13}$$

where O_X, O_Y are two operators supported respectively on two sets X, Y geometrically separated by a distance R, and $U_{0,T}^{\Lambda}$ is the time-evolution unitary of the full lattice Λ under a power-law decaying Hamiltonian, as defined above.

To compare different bounds, we analyze their effective light cones, which, up to constant prefactors, predict the minimum time it takes for the correlator $\mathcal{C}(T, R)$ to reach a certain value. For example, the original Lieb-Robinson bound [12] produces a linear light cone $T \gtrsim R$ for short-range interactions. For long-range interactions, Hastings and Koma [18] first showed that $\mathcal{C}(T, R) \leq ce^{vT}/R^{\alpha}$ for some (α-dependent) constants c, v. By setting $\mathcal{C}(T, R)$ equal to a constant, the bound gives an effective light cone $T \gtrsim \log R$ in the limit of large T and R. Gong et al. [29] later achieved a tighter light cone that is linear for short distances and becomes logarithmic only for large R. Shortly after, Foss-Feig et al. [19] derived a bound with a polynomial light cone:

$$T \gtrsim R^{\frac{\alpha-2D}{\alpha-D+1}}. \tag{5.14}$$

Equation (5.14) was the tightest light cone known previously.

In the remainder of this section, we use Lemma 3 to derive a Lieb-Robinson bound for long-range interactions that produces an effective light cone tighter than the one in Ref. [19], while also using a much more intuitive approach. In addition, our bound works for all times, unlike the bound in Ref. [19], which applies only in the long-time limit.

Theorem 7 (Lieb-Robinson bound for long-range interactions)**.** *Suppose O_X is supported*

on a fixed subset X. For $\alpha > 2D$, we have

$$\mathcal{C}(T,R) \leq \begin{cases} c_{\text{lr}} e^{\alpha} T R^{D-1} \xi_{\alpha}\left(\frac{R\alpha}{vT}\right), & \text{if } vT \geq \alpha, \\ \tilde{c}_{\text{lr}}\left(e^{vT}-1\right)\xi_{\alpha}(R), & \text{if } vT < \alpha. \end{cases} \tag{5.15}$$

Here $R = \text{dist}(X,Y)$ is the distance between the supports of O_X and O_Y, $c_{\text{lr}}, \tilde{c}_{\text{lr}}, v$ are constants that may depend only on D [defined in Appendix B.3], and ξ_{α} is given by Eq. (5.3).

Before we prove Theorem 7, let us analyze the features of the bound. Although the general bound in Eq. (5.15) looks complicated, it can be greatly simplified in some limits of interest. For example, for finite α, in the limit of large $vT > \alpha$ and large R such that $R/(vT) \gg \alpha$, the term algebraically decaying with $R/(vT)$ in $\xi_{\alpha}(R\alpha/(vT))$ dominates the exponentially decaying one [see also Eq. (5.3) and Eq. (B.47)]. Therefore, the Lieb-Robinson bound in this limit takes the form:

$$\mathcal{C}(T,R) \leq c_{\text{lr},\alpha} \frac{T^{\alpha-D}}{R^{\alpha-2D}}, \tag{5.16}$$

where $c_{\text{lr},\alpha}$ is finite and may depend on α [Eq. (B.49)]. We can immediately deduce the effective light cone given by our bound for a finite α:

$$T \gtrsim R^{\frac{\alpha-2D}{\alpha-D}}, \tag{5.17}$$

which is tighter than Eq. (5.14) (as given by Ref. [19]). In particular, for α close to $2D$, the exponent in Eq. (5.17) can be almost twice that of Ref. [19] (the larger the exponent,

the tighter the light cone).

On the other hand, in the limit $\alpha \to \infty$, vT is finite and therefore always less than α. Hence our bound converges to the short-range bound $\mathcal{C}(T, R) \leq 2\tilde{c}_{lr} e^{vT-\gamma R}$. We note that in this limit, the exponent of the light cone in Eq. (5.17) also converges to one, which corresponds to a linear light cone, at a linear convergence rate [see Eq. (B.51) for details]. These behaviors are naturally expected since a power-law decaying interaction with very large α is essentially a short-range interaction.

As mentioned earlier, we derive Theorem 7 by constructing a unitary $\tilde{U}$ such that *(i)* $\tilde{U}^\dagger O_X \tilde{U}$ approximates $\left(U_{0,T}^\Lambda\right)^\dagger O_X U_{0,T}^\Lambda$ and *(ii)* $\tilde{U}$ commutes with O_Y. We note that $\tilde{U}$ does not necessarily approximate $U_{0,T}^\Lambda$. It then follows from the two requirements that the commutator norm $\mathcal{C}(T, R)$, defined in Eq. (5.13), is upper bounded by the error of the approximation in (i).

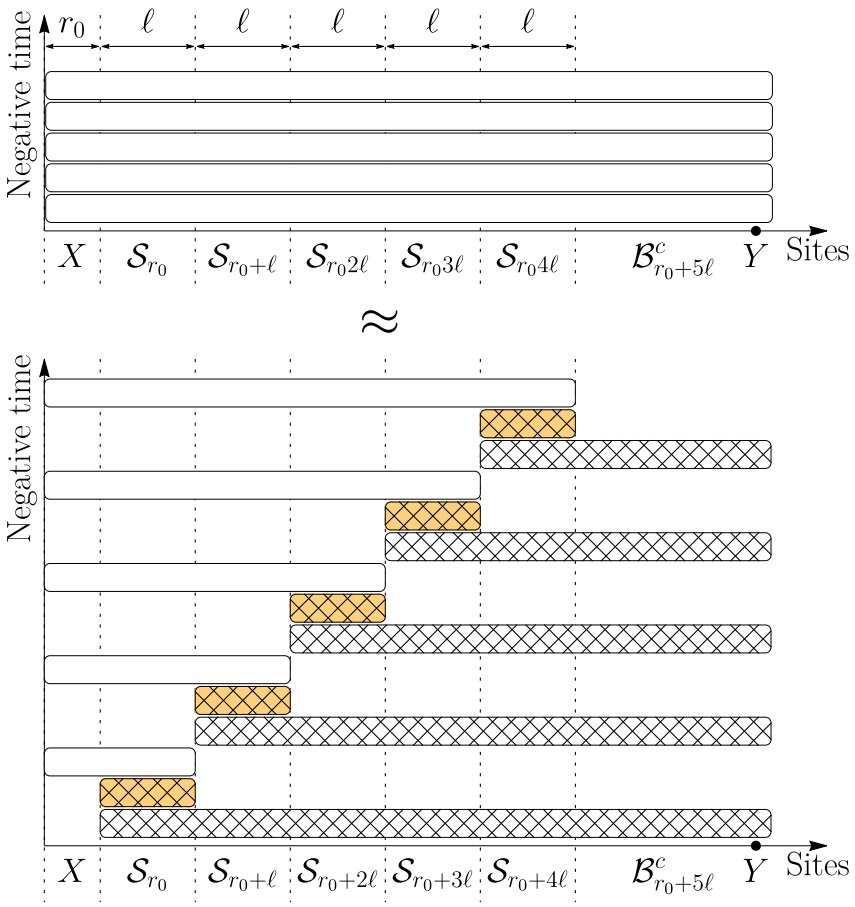

Figure 5.3: A construction of the unitary $\tilde{U}$ which results in an improved Lieb-Robinson bound for long-range interactions in Theorem 7. The horizontal axes list the sites in each subset. Here $\mathcal{B}_r$ denotes a D-ball of radius r centered on X, and $\mathcal{S}_r = \mathcal{B}_{r+\ell} \setminus \mathcal{B}_r$ a D-dimensional shell of inner radius r and outer radius $r + \ell$, for some parameter ℓ to be chosen later. (See Fig. B.2 in Appendix B.3 for an illustration of the sets.) The evolution unitaries are represented by boxes with the same color convention as in Fig. 5.1. We first divide the interval $[0, T]$ into $M = 5$ equal time slices (upper panel). Note that because we consider $O_X(T)$ in the Heisenberg picture, the vertical axis is therefore backward in time so that the bottom time slice will correspond to the first unitary applied on O_X. The evolution during each time slice is approximated by three evolutions of subsystems using Lemma 3 (lower panel). The bottom two unitaries have their supports outside X and therefore commute with O_X. They cancel with their Hermitian conjugates from $\tilde{U}^\dagger$ in $\tilde{U}^\dagger O_X \tilde{U}$. Repeating the argument for higher time slices, we can eliminate some unitaries (hatched boxes) from the construction of $\tilde{U}$. Finally, we are left with $\tilde{U}$ consisting only of unitaries (white boxes) that are supported entirely on the D-ball $\mathcal{B}_{r_0+5\ell}$ of radius $r_0 + 5\ell$. Therefore, $\tilde{U}$ commutes with O_Y, whose support lies in the complement $\mathcal{B}^c_{r_0+5\ell}$ of $\mathcal{B}_{r_0+5\ell}$.

We also note that the assumption on the norms of the interactions being bounded ex-

cludes several physical systems whose local dimensions are unbounded, e.g. bosons [see Ref. [119, 120] for discussions of information propagation and Lieb-Robinson bounds in these systems]. However, our Lieb-Robinson bound may still apply if the dynamics of the systems can be restricted to local Hilbert subspaces which are finite-dimensional. Examples of such situations include trapped ions in the perturbative regime [49] and non-interacting bosons [16].

To construct $\tilde{U}$, we use Lemma 3 to decompose the unitary $U_{0,T}^{\Lambda}$ into unitaries supported on subsystems, each of which either contains X or is disjoint from X. The unitaries of the latter type can be commuted through O_X to cancel out with their Hermitian conjugates from $\left(U_{0,T}^{\Lambda}\right)^{\dagger}$. The remaining unitaries form $\tilde{U}$, which is supported on a smaller subset than $U_{0,T}^{\Lambda}$. In particular, with a suitable decomposition, the support of $\tilde{U}$ can be made to not contain Y, and, therefore, $\tilde{U}$ commutes with O_Y. The step-by-step construction of the unitary $\tilde{U}$ has also been briefly described earlier in Sec. 5.1 and in Fig. 5.2, using the specific case of a one-dimensional system with a finite α. This construction immediately generalizes to higher dimensions and to arbitrary α, including the $\alpha \to \infty$ limit. The construction of $\tilde{U}$ for arbitrary D is summarized in Fig. 5.3.

We note that there is more than one way to decompose the unitary $U_{0,T}^{\Lambda}$ in the construction of $\tilde{U}$. Different constructions of $\tilde{U}$ result in different approximation errors, each of which provides a valid bound on the commutator norm $\mathcal{C}(T, R)$. Therefore, the goal is to find a construction of $\tilde{U}$ with the least approximation error. In Appendix B.3, we present the construction that results in the bound in Theorem 7. Although we have evidence suggesting that the construction is optimal, we do not rule out the existence of a better construction.

5.4 Faster Digital Quantum Simulation

In this section, we generalize the algorithm in Ref. [76] to simulating long-range interactions. In general, the aim of quantum simulation algorithms is to approximate the time evolution unitary $U_{0,T}^{\Lambda}$ using the fewest number of primitive, e.g. two-qubit, quantum gates. Here, we show that in addition to the stronger Lieb-Robinson bound presented in the previous section, Lemma 3 can also be used to perform error analysis for the HHKL algorithm (Ref. [76]) in the case of interactions that decay as a power law, therefore improving the theoretical gate count of digital quantum simulation for such interactions.

Using the best known rigorous error bounds, simulations based on the first-order Suzuki-Trotter product formula [1] use $\mathcal{O}\left(T^2 n^6/\varepsilon\right)$ gates to simulate the evolution $U_{0,T}^{\Lambda}$ of a time-dependent Hamiltonian on n sites up to a fixed error ε. (In this section, the big $\mathcal{O}$ is with respect to n, T, and $1/\varepsilon$.) The generalized $(2k)$th-order product formula uses $\mathcal{O}\left(n^2 (Tn^2)^{1+1/(2k)}/\varepsilon^{1/2k}\right)$ quantum gates. While this scaling asymptotically approaches $\mathcal{O}\left(Tn^4\right)$ as $k \to \infty$, it suffers from an exponential prefactor of 5^{2k} [6]. More advanced algorithms, e.g., those using quantum signal processing (QSP) [10] or linear combinations of unitaries (LCU) [8], can reduce the gate complexity to $\mathcal{O}\left(Tn^3 \log(Tn/\varepsilon)\right)$. Our error analysis below (Lemma 4) reveals that, when α is large, the number of quantum gates required by the HHKL algorithm to simulate long-range interactions scales better as a function of the system size than previous algorithms.

The HHKL algorithm itself uses either the QSP algorithm or the LCU algorithm as a subroutine to simulate the dynamics of a subset of the sites for one time step. Although the QSP algorithm does not handle time-dependent Hamiltonians, LCU can be applied

to time-dependent Hamiltonians. Our results assume that *(i)* the local terms $h_{\vec{i},\vec{i}}(t)$ have bounded norms for all $\vec{i} \in \Lambda$, and *(ii)* the Hamiltonian $H_\Lambda(t)$ varies slowly and smoothly with time so that $h'_{|X|} \equiv \max_t \left\| \partial H_t^X / \partial t \right\|$ exists and scales at most polynomially with $|X|$ for all subsets $X \subset \Lambda$. These restrictions allow portions of the system to be faithfully simulated using LCU (or QSP, for a time-independent Hamiltonian).

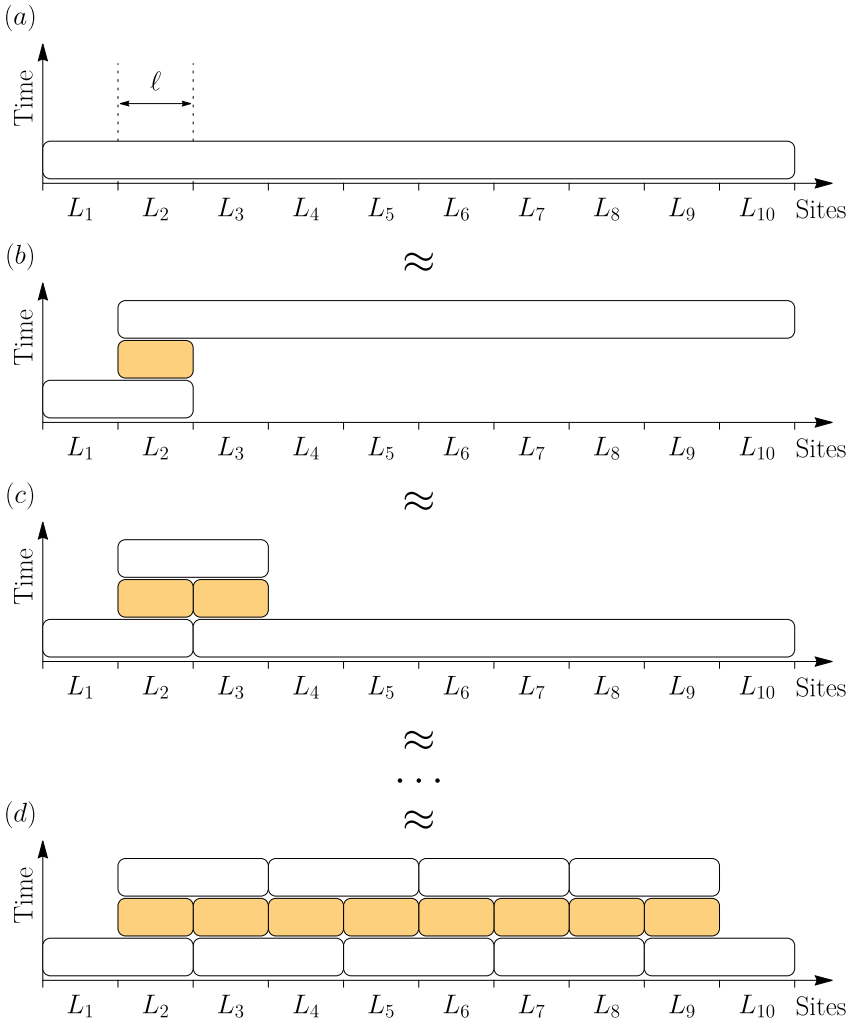

Figure 5.4: A demonstration of the HHKL decomposition [76] of the evolution of a fixed time interval for a system with $m = 10$ blocks, each consisting of ℓ sites. As before, each box represents a unitary (white) or its Hermitian conjugate (orange) supported on the covered sites. Using Lemma 3, the HHKL decomposition approximates the evolution of the whole system [panel (a)] by three unitaries supported on subsystems [panel (b)]. By applying Lemma 3 repeatedly [panels (c) and (d)], the evolution of the whole system is decomposed into a series of evolutions of subsystems, each of size at most 2ℓ.

5.4.1 HHKL-Type Algorithm for Simulating Long-Range Interactions

Although Ref. [76] focused on simulating short-range interactions, their (HHKL) algorithm can also be used to simulate long-range interactions. Here, we analyze the performance of their algorithm in simulating such systems. In the HHKL algorithm [76], the evolution of the whole system is decomposed, using Lemma 3, into *elementary* unitaries, each evolving a subsystem of at most $(2\ell)^D$ sites, where ℓ is again a length scale to be chosen later. For a fixed time t, the algorithm simply simulates each of these elementary unitaries using one of the existing quantum simulation algorithms. In particular, we shall use LCU or (for a time-independent Hamiltonian) QSP due to their logarithmic dependence on the accuracy.

In this section, we assume α is finite and analyze the gate count in the limit of large system size $n \gg \alpha$. As a consequence, the block size ℓ can also be taken to be much larger than α. For simplicity, we will not keep track of constants that may depend on α. Recall that in this limit, the error bound in Lemma 3 is at most

$$\mathcal{O}\left(\frac{\Phi(A)}{\ell^{\alpha-D-1}}\right), \tag{5.18}$$

where we have assumed $t = \mathcal{O}(1)$. Using Lemma 3, we obtain the error bound for the first step of the HHKL algorithm, which can be summarized by the following lemma.

Lemma 4 (HHKL decomposition)**.** *There exists a circuit that approximates $U_{0,T}^{\Lambda}$ up to error $\mathcal{O}\left(Tn/\ell^{\alpha-D}\right)$, where $\ell \leq n^{1/D}/2$ is a free parameter. The circuit has depth at most $3^D T$ and consists of $\mathcal{O}\left(Tn/\ell^D\right)$* elementary *unitaries, each of which evolves a subsystem*

supported on at most $(2\ell)^D$ *sites for time* $t = \mathcal{O}(1)$.

Proof. We now demonstrate the proof by constructing the circuit for a one-dimensional lattice (Fig. 5.4). A generalization of the proof to arbitrary dimension follows the same lines and is presented in Appendix B.4.

First, we consider $M \propto T$ equal time intervals $0 = t_0 < t_1 < \cdots < t_M = T$ such that $t_{j+1} - t_j = t = T/M$ is a constant for all $j = 0, \ldots, M-1$. The simulation of $U_{0,T}^{\Lambda}$ then naturally decomposes into M consecutive simulations of $U_{t_j,t_{j+1}}^{\Lambda}$. We then divide the system into m consecutive disjoint blocks, each of size $\ell = n/m$ (Fig. 5.4). Denote by L_k $(k = 1, \ldots, m)$ the set of sites in the k-th block. Using Lemma 3, we can approximate

$$U_{0,t}^{\Lambda} \approx U_{0,t}^{L_1 \cup L_2} \left(U_{0,t}^{L_2}\right)^{\dagger} U_{0,t}^{L_2 \cup L_3 \cup \cdots \cup L_m}. \tag{5.19}$$

This approximation can be visualized using the top two panels of Fig. 5.4. Repeated application of Lemma 3 yields the desired circuit (bottom panel of Fig. 4), with each elementary unitary evolving at most 2ℓ sites for time t.

To obtain the error estimate in Lemma 4, we count the number of times Lemma 3 is used in our approximation. In each of the M time slices, we use the lemma $\mathcal{O}(m) = \mathcal{O}(n/\ell)$ times, each of which contributes an error of $\mathcal{O}(1/\ell^{\alpha-2})$ [see Eq. (5.18) with $\Phi = \mathcal{O}(1)$ in one dimension]. Therefore, with $M \propto T$, the error of using the constructed circuit to simulate $U_{0,T}^{\Lambda}$ is

$$\mathcal{O}\left(M \frac{n}{\ell} \frac{1}{\ell^{\alpha-2}}\right) = \mathcal{O}\left(\frac{Tn}{\ell^{\alpha-1}}\right), \tag{5.20}$$

as given in Lemma 4. □

The error bound for the approximation in Lemma 4 leads to an upper bound on the gate complexity of digital quantum simulation, as stated in the following theorem.

Theorem 8. *For $\alpha > 2D$, there exists a quantum algorithm for simulating $U_{0,T}^{\Lambda}$ up to error at most ε with gate complexity*

$$G_D = \mathcal{O}\left(Tn\left(\frac{Tn}{\varepsilon}\right)^{\frac{2D}{\alpha-D}} \log\frac{Tn}{\varepsilon}\right). \tag{5.21}$$

This gate complexity can be achieved by applying the HHKL algorithm [76] for long-range interactions, as described above. First, the evolution of the whole system $U_{0,T}^{\Lambda}$ is approximated by $\mathcal{O}\left(Tn/\ell^D\right)$ elementary unitaries as provided in Lemma 4. Each of these elementary unitaries is then simulated using one of the existing algorithms, e.g., LCU, with error that we require to be at most $\varepsilon\ell^D/Tn$. If the Hamiltonian is time-independent, one can also use the QSP algorithm to simulate the elementary unitaries.

In the decomposition of the evolution, the accuracy of the approximation can be improved by increasing the block size ℓ. By Lemma 4, to achieve an overall error at most ε, we need

$$\ell \propto \left(\frac{Tn}{\varepsilon}\right)^{\frac{1}{\alpha-D}}. \tag{5.22}$$

When simulating the elementary unitaries, since each is an evolution of at most $(2\ell)^D$ sites for time $t = \mathcal{O}(1)$, the LCU algorithm with error at most $\varepsilon\ell^D/Tn$ uses $\mathcal{O}\left(\ell^{3D}\log\left(\frac{Tn}{\varepsilon}h'_{\ell^D}\right)\right)$ two-qubit gates [6]. Recall that we assume h'_{ℓ^D} scales at most polynomially with ℓ^D.

With the block size ℓ from Eq. (5.22), we find the total gate complexity of simulating the $\mathcal{O}\left(Tn/\ell^D\right)$ elementary unitaries is

$$G_D = \mathcal{O}\left(\frac{Tn}{\ell^D}\,\ell^{3D}\log\left(\frac{Tn}{\varepsilon}h'_{\ell^D}\right)\right) \tag{5.23}$$

$$= \mathcal{O}\left(Tn\left(\frac{Tn}{\varepsilon}\right)^{\frac{2D}{\alpha-D}}\log\frac{Tn}{\varepsilon}\right). \tag{5.24}$$

The scaling of G_D as a function of the system size n is significantly better than existing algorithms for large α. In particular, at $T = n$, this HHKL algorithm for long-range interactions requires only $\mathcal{O}\left(n^{2+\frac{4D}{\alpha-D}}\log n\right)$ gates, while algorithms such as QSP or LCU use $\mathcal{O}\left(n^4\log n\right)$ gates or more. Therefore, the algorithm provides an improvement for $\alpha > 3D$. However, the gate complexity of the algorithm depends polynomially on $1/\varepsilon$, in contrast to the logarithmic dependence achieved by QSP and LCU, and by the HHKL algorithm for systems with short-range interactions. While this poly$(1/\varepsilon)$ scaling is undesirable, in practice, the total error of the simulation is often set to a fixed constant (for example, see Ref. [121]) and effectively the dependence of ε only contributes a prefactor to the gate complexity of the algorithm.

As an example, in Fig. 5.5, we estimate the actual gate count of the HHKL algorithm in simulating a Heisenberg chain [Eq. (7.18)] and compare it with the gate count of the QSP algorithm (up to the same error tolerance). Because of the poly$(1/\varepsilon)$ overhead, the HHKL algorithm based on Lieb-Robinson bounds uses more quantum gates for simulating small systems, but eventually outperforms the QSP algorithm when the system size n is large.

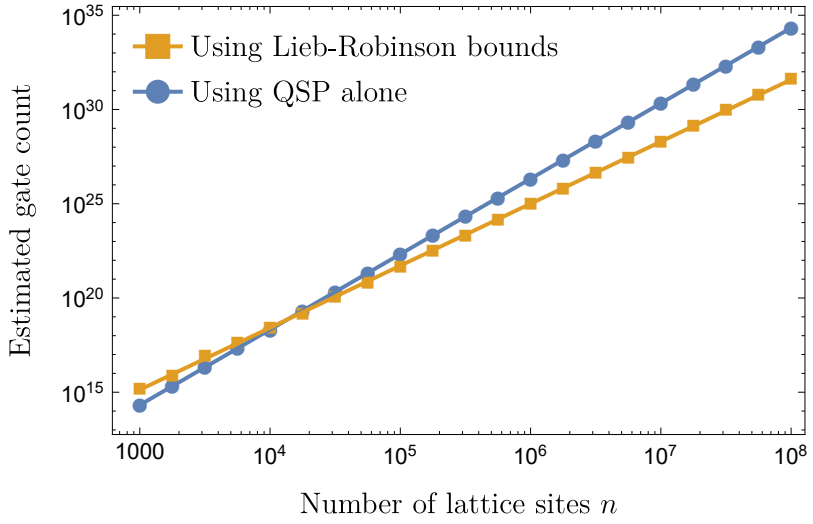

Figure 5.5: The gate count for simulating the dynamics of a one-dimensional Heisenberg chain [Eq. (7.18)] of length n, with $\alpha = 4, T = n$, and $\varepsilon = 10^{-3}$. We compare the gate count of the HHKL algorithm (orange square) to the QSP algorithm (blue circle). Note that the HHKL algorithm based on Lieb-Robinson bounds also uses the QSP algorithm as a subroutine. We obtain the scatter points using the approach described in Appendix B.5 and fit them to a power-law model (solid lines). The asymptotic scalings of the gate count obtained from the power-law fits ($n^{3.29}$ for HHKL, $n^{4.00}$ for QSP) agree well with our theoretical predictions (see Table 5.1).

It is also worth noting that, in the limit $\alpha \to \infty$, the gate complexity becomes $\mathcal{O}\left(Tn \log\left(Tn/\varepsilon\right)\right)$, which coincides (up to a polylogarithmic factor) with the result for short-range interactions in Ref. [76]. This behavior is expected, given that a power-law decaying interaction with $\alpha \to \infty$ is essentially a nearest-neighbor interaction. However, we caution readers that at the beginning of this section, we have assumed that α is finite so that $n \gg \alpha$. Hence, the gate count in Eq. (5.24) is technically not valid in the limit $\alpha \to \infty$. Nevertheless, the error bound in Lemma 3 reproduces the estimate for short-range interactions in Ref. [76], and therefore, repeating the argument of this section in the limit $\alpha \to \infty$ should also reproduce the gate count for simulating short-range interactions in Ref. [76].

5.4.2 Numerical Evidence of Potential Improvement

Up to now, we have seen that Lieb-Robinson bounds can improve the error bounds of quantum simulation algorithms, as demonstrated by the HHKL algorithm. We now provide numerical evidence hinting at the possibility of further improving the error bounds.

Although the HHKL algorithm outperforms previous ones when $\alpha > 3D$, it remains an open question whether there is a faster algorithm for simulating long-range interactions. We also note that the gate complexities are only theoretical upper bounds, and these algorithms may actually perform better in practice [7].

As an example, we compute the empirical gate count of a Suzuki-Trotter product formula simulation of a one-dimensional long-range interacting Heisenberg model

$$H = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{|i-j|^4} \vec{\sigma}_i \cdot \vec{\sigma}_j + \sum_{i=1}^{n} B_i \sigma_i^z, \tag{5.25}$$

where $B_j \in [-1, 1]$ are chosen uniformly at random and $\vec{\sigma}_j = (\sigma_j^x, \sigma_j^y, \sigma_j^z)$ denotes the vector of Pauli matrices on the qubit j. Specifically, we consider a simulation using the fourth-order product formula (PF4). We use a classical simulation to determine the algorithm's performance for systems of size $n = 4$ to $n = 12$ for time $T = n$, and extrapolate to larger systems. For each n, we search for the minimum number of gates for which the simulation error is at most $\varepsilon = 10^{-3}$. We plot in Appendix B.6 this empirical gate count, which appears to scale only as $\mathcal{O}\left(n^{3.64}\right)$ with the system size n. We list in Table 5.1 the gate counts of several popular algorithms for comparison. The *theoretical* gate complexity of PF4 is $\mathcal{O}\left(n^{5.75}\right)$ [1], while the QSP and LCU algorithms both have

Algorithm	Scaling with $n = T$	Scaling with ε
Empirical PF4	$\mathcal{O}\left(n^{3.64}\right)$	—
Our HHKL bound	$\mathcal{O}\left(n^{3.33}\log n\right)$	$\mathcal{O}\left(\log(1/\varepsilon)/\varepsilon^{0.67}\right)$
PF4 bound [6]	$\mathcal{O}\left(n^{5.75}\right)$	$\mathcal{O}\left(1/\varepsilon^{0.25}\right)$
QSP bound [10]	$\mathcal{O}\left(n^{4}\log n\right)$	$\mathcal{O}\left(\log(1/\varepsilon)\right)$
LCU bound [8]	$\mathcal{O}\left(n^{4}\log n\right)$	$\mathcal{O}\left(\log(1/\varepsilon)\right)$

Table 5.1: A comparison between the gate complexity of several quantum simulation algorithms for simulating one-dimensional power-law systems at $T = n$ and $\alpha = 4$. Our analysis shows that the HHKL algorithm performs at least as well as the empirical gate count of PF4, while having a similar poly$(1/\varepsilon)$ scaling with the error ε. It is not known whether the empirical gate count of PF4 can scale with ε better than suggested by the best proven bound (the third row).

complexity $\mathcal{O}\left(n^4 \log n\right)$. These numerics show that the PF4 algorithm for simulating long-range interacting systems performs better in practice than theoretically estimated; in fact, it even performs almost as well as the HHKL algorithm based on Lieb-Robinson bounds [which scales as $\mathcal{O}\left(n^{3.33}\log n\right)$ by our earlier analysis]. Whether other quantum simulation algorithms, including the HHKL algorithm, can perform better than suggested by the existing bounds remains an important open question.

5.5 Discussion

To conclude, we derived an improved bound on how quickly quantum information propagates in systems evolving under long-range interactions. The bound applies to power-law interactions with $\alpha > 2D$, such as dipole-dipole interactions in 1D (often realizable with nitrogen-vacancy centers [58] or polar molecules [59]), trapped ions in 1D [49, 50], and van-der-Waals-type interactions between Rydberg atoms [55] in either 1D or 2D. For finite $\alpha > 2D$, our Lieb-Robinson bound gives a tighter light cone than previously known bounds—including the one used in the proof of Lemma 3. As of yet,

we are not aware of any physical systems that saturate the Lieb-Robinson bounds for power-law interactions, including the new bound. In the limit $\alpha \to \infty$, our bound asymptotically approaches the exponentially decaying bound for short-range interactions. Our bound gives a linear light cone only in this limit, however, and it remains an open question whether there exists a stronger bound with a critical α_c such that the light cone is exactly linear for $\alpha \geq \alpha_c$ [78]. Currently, there are no known methods for quantum information transfer that are faster than linear for $\alpha \geq D + 1$. It is possible, therefore, that a stronger bound exists with a finite $\alpha_c \geq D + 1$. It is our hope that the present work, as well as the techniques that we use, will help motivate the search for such stronger bounds.

Our technique immediately extends the HHKL algorithm in Ref. [76] to the digital quantum simulation of the above systems. Our error bounds indicate that the gate complexity of the algorithm is better than that of other state-of-the-art simulation algorithms when α is sufficiently large ($\alpha > 3D$), and matches that of the short-range algorithm when $\alpha \to \infty$.

However, the empirical scaling of other algorithms—such as product formulas—indicates that this gate complexity may only be a loose upper bound to the true quantum complexity of the problem. While a matching lower bound for the gate complexity of the HHKL algorithm is provided in Ref. [76] for Hamiltonians with short-range interactions, we do not know of any techniques that could provide a corresponding bound for long-range interactions. In addition to improving the quantum gate complexity, our results may also aid in the design of better classical algorithms for simulating long-range interacting quantum systems. In particular, while we still expect the classical gate complexity to be exponential in the simulation time, there may be room for a polynomial improvement.

While the use of Lieb-Robinson bounds to improve the performance of quantum algorithms is a natural extension of Haah et al., the opposite direction—using quantum algorithms to improve Lieb-Robinson bounds—is new. In addition to proving a stronger Lieb-Robinson bound, the tools we developed may help to answer other questions regarding both short-range and long-range interacting systems. Using the same unitary construction as Theorem 7, we can generalize the bounds on connected correlators [36, 115] to long-range interacting systems. Our results can also provide a framework for proving tighter bounds on higher-order commutators, such as out-of-time-order correlators [116, 117]. Previous methods used to derive Lieb-Robinson bounds—due to their use of the triangle inequality early in their proofs—have not been able to capture the nuances in the growth of such correlators. In addition to the more intuitive proof of the Lieb-Robinson bounds, our framework can be used to provide an alternative, simpler proof of the classical complexity of the boson-sampling problem [82], which generalizes the result in Ref. [16] to long-range interactions and also to more general Hamiltonians with arbitrary local interactions [122]. By taking advantage of the unitary decomposition in Lemma 3, we obtain a longer time interval within which the sampler in Ref. [16] is efficient [122]. Moreover, by generalizing from two-body to many-body interactions, our technique may find applications in systems whose Hamiltonians include interaction terms between three or more sites, e.g. many-body localized systems in the l-bit basis [123].

Chapter 6: Other Implications of the Locality in Quantum Systems

In this Chapter, we discuss several implications of the locality of quantum systems besides the faster quantum simulation algorithms.

6.1 Growth of 2-Point Connected Correlators

In this section, we explore how fast connected correlators can be generated using a power-law Hamiltonian. In particular, we use the Lieb-Robinson bounds to show that the growth of connected correlators is constrained to linear light cones for all $\alpha > 2d + 1$. In contrast, for $\alpha < 2d + 1$, we construct—based on our protocol in Theorem 3—an explicit example where the growth of connected correlators is not contained within any linear light cone.

We consider a d-dimensional lattice Λ and a power-law Hamiltonian $H(t)$ with an exponent α. Let C denote a plane that separates Λ into two disjoint subsets L and R, with $L \cup R = \Lambda$. Let A and B be two unit-norm operators supported on single sites $x \in L$ and $y \in R$ respectively such that $\text{dist}(x, C), \text{dist}(y, C) > r/2$. Finally, let $|\psi\rangle$ be a product state between the sublattices L, R. Our object is the connected correlator

$$\mathcal{C}(t, r) = \langle A(t)B(t)\rangle - \langle A(t)\rangle \langle B(t)\rangle , \tag{6.1}$$

where $\langle\cdot\rangle = \langle\psi|\cdot|\psi\rangle$ and $A(t)$ is the time-evolved version of A under H. While the correlator vanishes at time zero due to the disjoint supports of A and B, it may grow with time as the operators spread under the evolution.

First, we show that $\mathcal{C}(t,r)$ obey a linear light cone for all $\alpha > 2d+1$. Our strategy is to approximate $A(t)$ by an operator $\tilde{A}$ supported on a ball of radius $r/2$ centered on x and $B(t)$ by $\tilde{B}$ supported on a ball of the same radius but centered on y. Since $\tilde{A}$ and $\tilde{B}$ have disjoint supports, the connected correlator between them vanishes. Therefore, the connected correlator between $A(t)$ and $B(t)$ is upper bounded by the errors of the approximations:

$$\mathcal{C}(t,r) \leq 2\|A(t) - \tilde{A}\| + 2\|B(t) - \tilde{B}\|, \tag{6.2}$$

which in turn depend on the constructions of $\tilde{A}, \tilde{B}$.

Let S_A contain sites that are at most a distance $r/2$ away from x and S_A^c be all other sites in the lattice. We construct $\tilde{A}$ by simply tracing out the part of $A(t)$ that lies outside S_A [36]:

$$\tilde{A} = \mathrm{tr}_{S_A^c}[A(t)], \tag{6.3}$$

where the partial trace is taken over S_A^c. It follows from the definition that $\tilde{A}$ is supported entirely on S_A. Proposition 2 provides a bound on the error in approximating $A(t)$ by $\tilde{A}$:

there exists a velocity u such that when $r > ut$,

$$\|A(t) - \tilde{A}\| \leq K \frac{t^{d+1} \log^{2d} r}{r^{\alpha - d}}, \tag{6.4}$$

for some constant $0 < K < \infty$. Upper bounding the error in approximating $B(t)$ by $\tilde{B}$ in a similar way, we obtain a bound on the connected correlator:

$$\mathcal{C}(t, r) \leq 4K \frac{t^{d+1} \log^{2d} r}{r^{\alpha - d}}. \tag{6.5}$$

As a result, the light cone of the connected correlator is linear, with velocity no larger than u, for $\alpha > 2d + 1$.

We now provide an example of superlinear growth of connected correlators for $\alpha < 2d + 1$ using a slightly altered protocol than that discussed in Section 4.3.2. In particular, we consider initial operators $A = X_x$ and $B = Z_y$ supported on x, y respectively such that dist$(x, y) = r$.

The protocol works as follows. In the first step of the protocol (again in time $t/3$), we still apply $\mathcal{U}_1$ in order to spread X_x to $\prod_{i \in \mathcal{B}_\ell} X_i$, where $\mathcal{B}_\ell$ is a ball of radius $\ell = t/3$ centered on x. Since $\mathcal{U}_1$ acts trivially on $\tilde{\mathcal{B}}_\ell$ (the ball of radius ℓ centered on y), we can simultaneously apply a locally rotated version of $\mathcal{U}_1$ in $\tilde{\mathcal{B}}_\ell$ to spread Z_y to $\prod_{\tilde{i} \in \tilde{\mathcal{B}}_\ell} Z_{\tilde{i}}$. In the next time $t/3$, we again apply $\mathcal{U}_2$, which takes $\prod_{i \in \mathcal{B}_\ell} X_i$ to the expression in Eq. (4.37). Note that this evolution does not change $\prod_{\tilde{i} \in \tilde{\mathcal{B}}_\ell} Z_{\tilde{i}}$ as it commutes with $\mathcal{U}_2$. For the last $t/3$ we simply do nothing.

Define the state $|\psi\rangle = |\phi\rangle_{\mathcal{B}_\ell} |\phi\rangle_{\tilde{\mathcal{B}}_\ell}$, where

$$|\phi\rangle_{\mathcal{B}_\ell} = \frac{1}{\sqrt{2}}\big(\underbrace{|0\cdots 0\rangle_{\mathcal{B}_\ell}}_{\equiv|\bar{0}\rangle_{\mathcal{B}_\ell}} + i \underbrace{|1\cdots 1\rangle_{\mathcal{B}_\ell}}_{\equiv|\bar{1}\rangle_{\mathcal{B}_\ell}} \big) \tag{6.6}$$

is a state of the sites in $\mathcal{B}_\ell$. Throughout our analysis, we will often dispense with the subscripts, but the Hilbert space in question should be clear from context, and we will always list the state on $\mathcal{B}_\ell$ before that for $\tilde{\mathcal{B}}_\ell$.

We will calculate the connected correlator

$$\mathcal{C}(t,r) = \langle Z_y(t) X_x(t)\rangle - \langle Z_y(t)\rangle \langle X_x(t)\rangle , \tag{6.7}$$

where $\langle\cdot\rangle = \langle\psi|\cdot|\psi\rangle$ and $X_x(t), Z_y(t)$ are the operators evolved under the unitaries described above. Assume for simplicity that t is such that $\mathcal{V}$—the volume of $\mathcal{B}_\ell$—is odd. It is straightforward to show that $\langle Z_y(t)\rangle = 0$ and therefore the second term $\mathcal{C}(t,r)$ vanishes. Next, we have

$$\begin{aligned}
&X_x(t)\,|\psi\rangle \\
&= \prod_{j\in\mathcal{B}_\ell} \left[\cos(2\tau\Theta) X_j + \sin(2\tau\Theta) Y_j\right] |\psi\rangle \\
&= \frac{1}{\sqrt{2}} \prod_{j\in\mathcal{B}_\ell} \left[cX_j + sY_j\right] |\phi\rangle\,|\bar{0}\rangle + \frac{i}{\sqrt{2}} \prod_{j\in\mathcal{B}_\ell} \left[cX_j - sY_j\right] |\phi\rangle\,|\bar{1}\rangle \\
&= \frac{1}{2}\Big[(c+is)^{\mathcal{V}} |\bar{1}\rangle\,|\bar{0}\rangle + i(c-is)^{\mathcal{V}} |\bar{0}\rangle\,|\bar{0}\rangle + i(c-is)^{\mathcal{V}} |\bar{1}\rangle\,|\bar{1}\rangle - (c+is)^{\mathcal{V}} |\bar{0}\rangle\,|\bar{1}\rangle \Big],
\end{aligned} \tag{6.8}$$

where $c = \cos(2\tau\mathcal{V})$ and $s = \sin(2\tau\mathcal{V})$. Next note that:

$$\langle\psi| Z_y(t) = \frac{1}{\sqrt{2}}(\langle\phi| \langle\bar{0}| - i \langle\phi| \langle\bar{1}|)Z_y(t) = \frac{1}{\sqrt{2}}(\langle\phi| \langle\bar{0}| + i \langle\phi| \langle\bar{1}|) \tag{6.9}$$

$$= \frac{1}{2}\left(\langle\bar{0}| \langle\bar{0}| + i \langle\bar{0}| \langle\bar{1}|\right) - i \langle\bar{1}| \langle\bar{0}| + \langle\bar{1}| \langle\bar{1}|\right). \tag{6.10}$$

Thus:

$$\mathcal{C}(t, r) = \langle Z_y(t) X_x(t)\rangle = \frac{i}{2}((c - is)^{\mathcal{V}} - (c + is)^{\mathcal{V}}) \geq \frac{1}{3^{1+2d}2^{2+\alpha}} \frac{t^{2d+1}}{r^{\alpha}}, \tag{6.11}$$

where we have used the bound Eq. (4.43). Therefore, this demonstrates the connected correlators may grow along a superlinear light cone for all $\alpha < 2d + 1$.

We note that in our setting, we only assume the initial state is a bipartite product state across the cut C. Our bound therefore also applies to a more restrictive scenario where the initial states are fully product. However, it is not clear whether the bound can be saturated in this scenario.

6.2 Growth of n-Point Connected Correlators

In this section, we generalize Lieb-Robinson bounds on bipartite connected correlators to multipartite connected correlators. We then show that there exist systems where the bounds are saturated. We argue that the bounds on multipartite correlations provide practical advantages over bipartite bounds. In addition, our Lieb-Robinson bounds on multipartite connected correlators imply that exponentially large correlations can be created in fixed time, independent of a system's size. We provide explicit examples of sys-

tems with this feature.

6.2.1 Connected correlations

Let us first define bipartite connected correlators. Consider a set of n sites Γ and two distinct, non-overlapping subsets $\mathcal{X} \subset \Gamma$ and $\mathcal{Y} \subset \Gamma$. Denote by $\mathcal{S}(\mathcal{X})$ the set of observables for which support lies entirely in $\mathcal{X}$. The bipartite *disconnected* correlator between observables $A_\mathcal{X} \in \mathcal{S}(\mathcal{X})$ and $A_\mathcal{Y} \in \mathcal{S}(\mathcal{Y})$ is simply the expectation value of their joint measurement outcomes at equal time, i.e., $\langle A_\mathcal{X} A_\mathcal{Y} \rangle$. Often in experiments only single sites are directly accessible. Observables are then supported by single sites, i.e., $|\mathcal{X}| = |\mathcal{Y}| = 1$. In the following discussions we refer to such correlators as two-point disconnected correlators.

We note that disconnected correlators contain both quantum and classical correlations. For example in two-qubit systems, the disconnected correlator $\langle Z_1 Z_2 \rangle$ (where Z is the Pauli matrix) achieves maximal value in both the fully classical state $|00\rangle$ and the maximally entangled state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ [124]. Their difference lies in the local expectation values $\langle Z_1 \rangle$, $\langle Z_2 \rangle$, which are maximal for the product state and vanish for the maximally entangled state. These local expectation values therefore can be said to carry classical information of the systems (in pure states). The bipartite *connected* correlator is constructed by subtracting this "classicalness" from the disconnected correlator:

$$u_2(A_\mathcal{X}, A_\mathcal{Y}) \equiv \langle A_\mathcal{X} A_\mathcal{Y} \rangle - \langle A_\mathcal{X} \rangle \langle A_\mathcal{Y} \rangle . \tag{6.12}$$

In general for mixed systems, if the joint state of $\mathcal{X} \cup \mathcal{Y}$ is a product state, i.e., $\rho_{\mathcal{X}\cup\mathcal{Y}} = \rho_{\mathcal{X}} \otimes \rho_{\mathcal{Y}}$, its disconnected correlators $\langle A_{\mathcal{X}} A_{\mathcal{Y}} \rangle$ are factorizable into $\langle A_{\mathcal{X}} \rangle \langle A_{\mathcal{Y}} \rangle$ and therefore all bipartite connected correlators vanish. The opposite is also true [125]:

Lemma 5. *A density matrix ρ is a product state, i.e., there exist complementary subsets $\mathcal{X}, \tilde{\mathcal{X}}$ such that $\rho = \rho_{\mathcal{X}} \otimes \rho_{\tilde{\mathcal{X}}}$, if and only if*

$$u_2(A_{\mathcal{X}}, A_{\tilde{\mathcal{X}}}) = 0, \tag{6.13}$$

for all observables $A_{\mathcal{X}} \in \mathcal{S}(\mathcal{X})$ and $A_{\tilde{\mathcal{X}}} \in \mathcal{S}(\tilde{\mathcal{X}})$.

In particular, a nonzero bipartite connected correlator implies bipartite entanglement in pure states. Lemma 5 is a consequence of Ref. [125]. We also present a simple proof in Appendix C.1.

A natural generalization of the bipartite connected correlator to multipartite systems is the Ursell function [126, 127]. The n-partite connected correlator between n observables $A_1, \ldots, A_n$, which are supported by n distinct subsets of sites $\mathcal{X}_1, \ldots, \mathcal{X}_n$, respectively, is defined as

$$u_n(A_1, \ldots, A_n) = \sum_P g(|P|) \prod_{p \in P} \left\langle \prod_{j \in p} A_j \right\rangle, \tag{6.14}$$

where $g(x) = (-1)^{x-1}(x-1)!$ and the sum is taken over all partitions P of the set $\{1, 2, \ldots, n\}$. The n-partite connected correlators can be equivalently defined via either recursive relations or generating functions (see Appendix C.2 for details).

Multipartite connected correlators also arise naturally in many other contexts. In

quantum field theory, connected Green's functions are multipartite connected correlators of field operators [128]. Mean field theory is an approximation in which it is assumed that all connected correlators vanish [129]; in fact, mean field theory fails when there exist significant connected correlations, and one must then seek higher-order approximations. The cumulant expansion technique is similar to mean field theory, but only multipartite connected correlators of high enough order are ignored. Therefore, understanding when connected correlations are negligible is important for validating mean field theory and the cumulant expansion.

The relation mentioned above between connected correlators and entanglement holds for n-partite connected correlators as well. It also follows from Ref. [125] that n-partite connected correlators vanish in product states. In particular, for pure states, a nonzero n-partite connected correlator implies genuine n-partite entanglement [130, 131]:

Lemma 6. *If an n-partite system is in a product state, i.e. there exist complementary subsystems $\mathcal{X}, \bar{\mathcal{X}} \subset S_n$ such that*

$$\rho = \rho_{\mathcal{X}} \otimes \rho_{\bar{\mathcal{X}}}, \tag{6.15}$$

then all k-body connected correlators ($2 \leq k \leq n$) between some observables $A_1, \ldots, A_{k_1}$, for which support lies entirely on $\mathcal{X}$, and observables $B_1, \ldots, B_{k_2}$, for which support lies entirely on $\bar{\mathcal{X}}$ ($k_1, k_2 \geq 1, k_1 + k_2 = k$), vanish,

$$u_k\left(A_1, \ldots, A_{k_1}, B_1, \ldots, B_{k_2}\right) = 0. \tag{6.16}$$

Corollary 1. *If an n-partite pure state $|\psi\rangle$ has a nonzero n-partite connected correlator, then it is genuinely n-partite entangled, i.e. there exist no subsystems $\mathcal{X}$ and $\tilde{\mathcal{X}}$ such that $|\psi\rangle = |\psi_{\mathcal{X}}\rangle \otimes |\psi_{\bar{\mathcal{X}}}\rangle$.*

A direct proof of Lemma 6 is presented in Appendix C.3. The combination of Lemma 1 and Lemma 2 tells us that if the bipartite connected correlators are all zero between two regions, then all higher-order connected correlators are guaranteed to be zero except for the scenario where all observables are supported on one region, or there exists an observable supported on both regions.

Multipartite connected correlations also provide a practical advantage over bipartite correlations, even though the latter are sufficient to characterize a quantum system. Consider a three-body system for example. The collection of local expectation values and connected correlators,

$$U = \Big\{ \langle A_1 \rangle , \langle A_2 \rangle, \langle A_3 \rangle , u_2(A_1, A_2), u_2(A_1, A_3), \\ u_2(A_2, A_3), u_3(A_1, A_2, A_3) \Big\}, \tag{6.17}$$

where each A_j runs over a complete single site basis (e.g. the Pauli matrices X, Y, Z), defines a unique tripartite quantum state. Another equivalent collection $\tilde{U}$ can be constructed from U by replacing $u_3(A_1, A_2, A_3)$ with a bipartite connected correlator between one subsystem and the rest, e.g. $u_2(A_1, A_2A_3)$. Although the two collections U and $\tilde{U}$ are equivalent, $u_3(A_1, A_2, A_3)$ and $u_2(A_1, A_2A_3)$ carry different information about the system. The three-point connected correlators $u_3(A_1, A_2, A_3)$ characterize global properties

while $u_2(A_1, A_2A_3)$ only tell us about local properties across the cut between subsystem 1 and the rest. If global properties, such as genuine three-body entanglement, are of concern, then tripartite connected correlators are superior. To have a chance at detecting genuine tripartite entanglement using only bipartite connected correlators, one must consider all possible bipartitions of the system. There are only 3 such partitions for a tripartite system, namely $1|23, 2|13$ and $3|12$. But for n-partite systems, the number of bipartitions scales exponentially with n. Computing all of them would be impractical. Even then there is no guarantee they would detect genuine multipartite entanglement. Consider for example the following pure state of 3 qubits,

$$\begin{aligned}|\psi\rangle =&\sqrt{\frac{5}{24}}\,|000\rangle + \sqrt{\frac{1}{8}}\,|001\rangle + \sqrt{\frac{1}{12}}\,|010\rangle + \sqrt{\frac{1}{12}}\,|011\rangle \\ &+ \sqrt{\frac{1}{4}}\,|100\rangle + \sqrt{\frac{1}{8}}\,|101\rangle + \sqrt{\frac{1}{12}}\,|110\rangle + \sqrt{\frac{1}{24}}\,|111\rangle\,. \end{aligned} \tag{6.18}$$

Its three-point connected correlator $u_3(Z_1, Z_2, Z_3) = \frac{1}{18}$ implies genuine tripartite entanglement in $|\psi\rangle$. Meanwhile, non-zero bipartite connected correlators across the cuts $2|13$ and $3|12$, $u_2(Z_2, Z_1Z_3)$ and $u_2(Z_3, Z_1Z_2)$, only tell us that there is entanglement between qubits 2 and 3. Because the bipartite connected correlator across $1|23$ $u_2(Z_1, Z_2Z_3)$ is zero, it is inconclusive whether the first qubit is entangled with the others without considering higher order correlators.

This example demonstrates why multipartite connected correlators are better candidates than bipartite counterparts in multipartite entanglement detection schemes. It is therefore important to understand how these multipartite correlations evolve in physical systems.

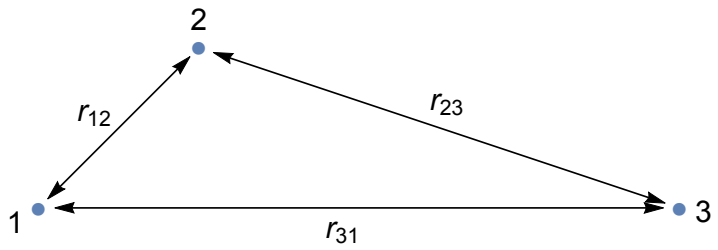

Figure 6.1: A typical three-body system. Each dot represents one site. There are three relevant length scales r_{12}, r_{23} and r_{31}. Which length scale will define the three-body Lieb-Robinson bound?

6.2.2 Multipartite Lieb-Robinson bounds

Our main result is Lieb-Robinson-like bounds on n-partite connected correlators in systems evolving from fully product states under short-range interactions, e.g.

$$H = \sum_{\langle i,j \rangle} J_{ij} V_i V_j, \tag{6.19}$$

where V_i is the spin operator of the i^{th} site, $|J_{ij}| \leq 1$ is the interaction strength between the i^{th} and the j^{th} sites, and the sum is over all neighboring i, j. But before we present the bounds, let us discuss general features we expect from such bounds. These bounds are of the form

$$u_n \leq C_n \exp(v_{\text{LR}} t - r), \tag{6.20}$$

where C_n is a constant, r is a relevant length scale, and v_{LR} is the same Lieb-Robinson velocity as in the bipartite bounds. Let us now examine the scaling of C_n with n. If all observables have unit norm, bipartite connected correlators are upper bounded by 1

regardless of a system's size. However, multipartite connected correlators can increase in value with the number of subsystems. For example, in the n-qubit Greenberger-Horne-Zeilinger (GHZ) state,

$$|\mathrm{GHZ}\rangle = \frac{|0\rangle^{\otimes n} + |1\rangle^{\otimes n}}{\sqrt{2}}. \tag{6.21}$$

the n-point connected correlator $u_n(Z_1, \ldots, Z_n) = \mathcal{O}(n^n)$ (details in Appendix C.5). Therefore we expect C_n to grow with n as well, $C_n = \mathcal{O}(n^n)$. Another constant we would like to understand is the critical distance r. In the Lieb-Robinson bound on a bipartite connected correlator, the critical distance is simply the distance between the two involved parties. However, in a multipartite system there are many relevant length scales which could possibly serve as the critical distance. As an example, let us consider a three-qubit system (Fig. 6.1). Without loss of generality we assume $r_{12} < r_{23} < r_{31}$ where r_{ij} denotes the distance between the i^{th} and j^{th} qubits. We argue that a bound of the form (6.20) with $r = r_{12}$ is valid but trivial. Intuitively an observable initially localized at the first qubit will need time to spread a distance r_{12} before "seeing" another qubit. Is there a stronger bound, i.e. inequality (6.20) with larger value for r? The largest distance r_{31} would make the most sense, since at $t = r_{31}/v$, an observable initially localized at one qubit has enough time to spread to all others. We shall show below that the critical distance for the tightest bound is neither the smallest (r_{12}) nor the largest distance (r_{31}), but actually the intermediate length scale r_{23}. This surprising result leads to unexpected consequences, including the creation of exponentially large connected correlations in unit time.

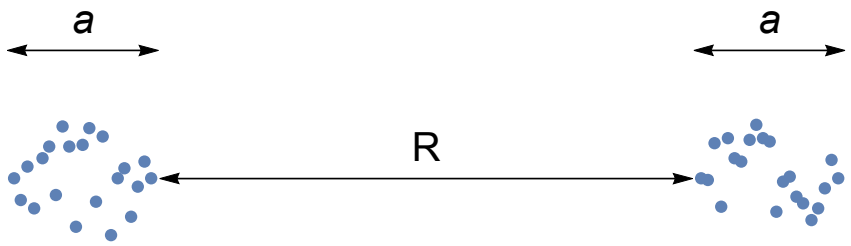

Figure 6.2: A geometry where n sites (blue dots) are divided into two cliques such that the clique size a is much smaller than the distance R between cliques.

Theorem 9. *Given n non-overlapping subsystems $\{\mathcal{X}_1, \ldots, \mathcal{X}_n\} = S$ initialized to a fully product state $|\psi_{\mathcal{X}_1}\rangle \otimes \cdots \otimes |\psi_{\mathcal{X}_n}\rangle$ and evolved under short-range interactions, the n-partite connected correlator between observables $A_i \in \mathcal{S}(\mathcal{X}_i)$ $(i = 1, \ldots, n)$ is bounded,*

$$|u_n(A_1, \ldots, A_n)| \leq C_n \exp(v_{\mathrm{LR}} t - R), \tag{6.22}$$

where v_{LR} is the same velocity as in the bipartite Lieb-Robinson bounds, $C_n = \frac{n^n}{4} C_2$ with C_2 being the constant in bipartite Lieb-Robinson bounds [1], and

$$R = \max_{\mathcal{S}_1 \subset S} d(\mathcal{S}_1, \bar{\mathcal{S}}_1) \tag{6.23}$$

is the largest distance between any subset $\mathcal{S}_1 \subset \mathcal{S}$ and its complementary subset $\bar{\mathcal{S}}_1$. Here the distance d between two sets of sites is the shortest distance between a site in one set and a site in the other set.

Proof. We shall explain our proof in the simplest case of $n = 3$. We use the following identity (given in Appendix C.2) to write disconnected correlators in terms of connected

[1] See Eq. (6) of Ref. [88]

correlators,

$$\begin{aligned}\langle A_1 A_2 A_3\rangle =& u_3(A_1,A_2,A_3) + u_2(A_2,A_3)\langle A_1\rangle \\ &+ u_2(A_1,A_3)\langle A_2\rangle + u_2(A_1,A_2)\langle A_3\rangle \\ &+ \langle A_1\rangle\langle A_2\rangle\langle A_3\rangle . \end{aligned} \tag{6.24}$$

Notice that the last two terms on the right hand side sum up to $\langle A_1 A_2\rangle\langle A_3\rangle$. If we move this term to the left hand side, we obtain an expression of u_3 in terms of only bipartite connected correlators (and local expectation values),

$$\begin{aligned} u_3(A_1,A_2,A_3) =& u_2(A_1A_2,A_3) - u_2(A_1,A_3)\langle A_2\rangle \\ &- u_2(A_2,A_3)\langle A_3\rangle , \end{aligned} \tag{6.25}$$

where the local expectation values $\langle A_2\rangle, \langle A_3\rangle$ are between -1 and 1. Therefore we may bound the three-body connected correlator using the bipartite Lieb-Robinson bound as follows,

$$\begin{aligned} &|u_3(A_1,A_2,A_3)| \\ &\leq |u_2(A_1A_2,A_3)| + |u_2(A_1,A_3)| + |u_2(A_2,A_3)| \\ &\leq C_2 e^{v_{\mathrm{LR}}t - r_{12|3}} + C_2 e^{v_{\mathrm{LR}}t - r_{13}} + C_2 e^{v_{\mathrm{LR}}t - r_{23}} \\ &\leq 3C_2 e^{v_{\mathrm{LR}}t - r_{12|3}}, \end{aligned} \tag{6.26}$$

where $r_{12|3} = \min\{r_{12}, r_{13}\}$ is the distance from the third site to the other two and C_2

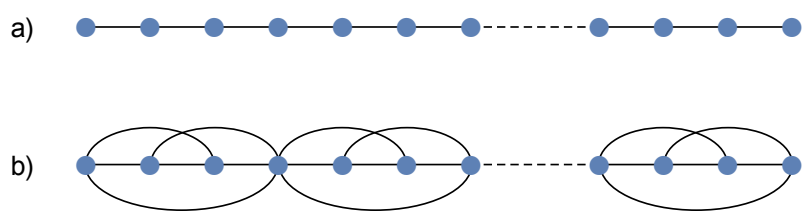

Figure 6.3: n-qubit cluster states represented by one-dimensional graphs of n vertices. (a) Only consecutive vertices are connected by edges of length 1. (b) Some edges are longer than 1 but interactions are still local.

comes from bipartite Lieb-Robinson bounds [88]. One may notice that at the beginning the three sites play equal roles, but somehow this symmetry is broken in Eq. (6.26). The reason is the choice to team up $\langle A_1\rangle \langle A_2\rangle \langle A_3\rangle$ and $u_2(A_1, A_2) \langle A_3\rangle$ after Eq. (6.24). Instead, we may replace the latter with either $u_2(A_2, A_3) \langle A_1\rangle$ or $u_2(A_1, A_3) \langle A_2\rangle$ to obtain two different bounds in the form of Eq. (6.26), with either $r_{23|1}$ or $r_{13|2}$ in place of $r_{12|3}$. The tightest bound corresponds to the smallest distance among $r_{23|1}, r_{13|2}, r_{12|3}$, and hence the theorem follows. Proof for general n follows the exact same line and is presented in full in Appendix C.4. □

Since the proof is inductive on the number of sites n, the multipartite Lieb-Robinson bounds are in general weaker than bipartite Lieb-Robinson bounds. Violation of our bound for a multipartite connected correlator implies violation of at least one bipartite bound. Nevertheless, the multipartite Lieb-Robinson bounds in Theorem 9 can be saturated. For example, consider a geometry of n sites where they are divided into two non-empty cliques, each of spatial size a. The two cliques are separated by a large distance $R \gg a$ (Fig. 6.2). Lieb-Robinson bounds of n-partite connected correlators for this geometry are saturated by preparing the GHZ state of n qubits, which can be done in time

$t \approx R/v_{\mathrm{LR}}$.

Whether our n-partite Lieb-Robinson bounds are tight for all geometries is still an open question. The geometry of Fig. 6.2 resembles a bipartite system, where each clique plays the role of one party. There are geometries which are very different from bipartite systems and, as a consequence, they offer some unique and interesting implications. For example, as mentioned before, the critical distance in the multipartite Lieb-Robinson bound is neither the largest nor the smallest distance. In the asymptotic limit of large n, these quantities can be very different. We shall now examine such examples.

6.2.3 Fast generation of multipartite correlation

In a bipartite system, the time needed to create bipartite correlators of order $\mathcal{O}(1)$ increases proportionally to the distance between the two subsystems. It is natural to expect the time needed to create n-point correlators of order $\mathcal{O}(1)$ in an n-partite system to increase with the spatial size of the system. However, Theorem 1 suggests that it may not necessarily be the case. For example, consider an equally spaced one-dimensional chain of n qubits [see Fig. 6.3]. If the distance between two consecutive qubits is fixed, the spatial length of the chain increases as $\mathcal{O}(n)$. Therefore 2-point connected correlators between the end qubits can only be created after $\mathcal{O}(n)$ time. Meanwhile, Theorem 1 suggests that n-point connected correlators of order $\mathcal{O}(1)$ between all n qubits might be created in $\mathcal{O}(1)$ time using only nearest neighbor interactions, enabling almost instant n-partite genuine entanglement. As we show below, systems with such a feature do exist.

One example is the one-dimensional cluster state. Cluster states (also called graph states) are multipartite entangled states [132] useful for one-way quantum computation [133, 134]. They have a simple visual representation using associated graphs. For a graph $G = (V, E)$, the corresponding cluster state can be constructed as follows: (i) Associate each vertex in V with one qubit initialized in state $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$; and (ii) Apply a controlled-Z gate to every pair of qubits connected by an edge in E. A controlled-Z gate on two qubits i and j can be implemented by evolving the system for a time $\frac{\pi}{4}$ under the Hamiltonian,

$$H_{cZ}^{(i,j)} = \mathbb{I} + Z_i + Z_j - Z_i Z_j, \tag{6.27}$$

where Z is the diagonal Pauli matrix. Some cluster states, e.g. Fig. 6.3, only require application of finite-range controlled-Zs. Meanwhile, the generating Hamiltonians (6.27) commute with each other and therefore they can be applied simultaneously. Therefore such cluster states as well as their correlations can be created in constant time $\mathcal{O}(1)$. In particular, within an n-independent time $\frac{\pi}{4}$ we can create $|u_n(Y_1, X_2, X_3 \ldots, X_{n-1}, Y_n)| = 1$ in a cluster state with only nearest neighbor interactions (Fig. 6.3(a). This example shows that n-point connected correlators of order $\mathcal{O}(1)$ can be created in unit time $\mathcal{O}(1)$, independent of a system's size. Yet, we can do better, i.e. we can create *exponentially large* n-point connected correlators in unit time. For example, in the cluster state of Fig. 6.3b, we allow each site to interact within a larger neighborhood. It still takes $\frac{3\pi}{4} = \mathcal{O}(1)$ unit time to prepare the state, while direct calculation shows that one of its correlators grows exponentially as $2^{\frac{n-1}{3}}$ (Appendix C.5).

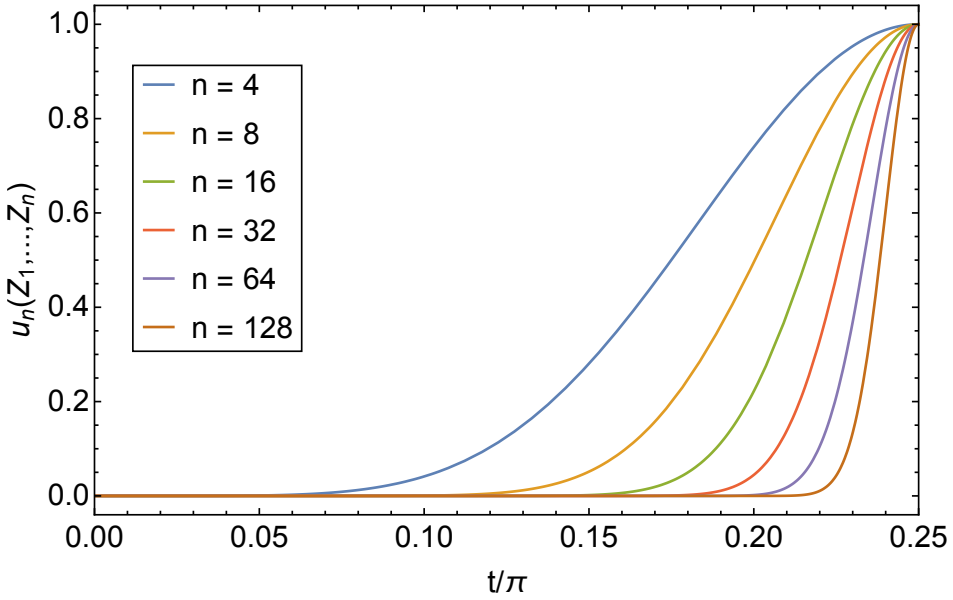

Figure 6.4: Time evolution of the n-point connected correlator $u_2(Z_1, Z_2, \ldots, Z_n) = \left[\sin^2(2t)\right]^{n-1}$ of the state in Fig. 6.3a for different n. Here we plot the time-dependent correlator for a few values of n. In the limit of large n the correlator remains zero for most of the time before briefly jumping to 1 at $t = \frac{\pi}{4}$.

In the above examples we have discussed how much time it takes to grow connected correlations from fully uncorrelated states. A relevant question is whether it can be expedited by some initial correlations [135]. To answer this question, we look at the time dependence of connected correlators in an n-qubit system initialized to $|00\ldots0\rangle$ and evolved under the Hamiltonian $\sum_{\langle i,j\rangle} X_i X_j$. If this system has the geometry of Fig. 6.3a, we find the n-point connected correlator $u_n(Z_1, \ldots, Z_n) = \left[\sin^2(2t)\right]^{n-1}$ (Appendix C.5). We plot this function for a few values of n in Fig. 6.4. For large n the correlator remains negligible for most of the time and rapidly grows to 1 near $t = \frac{\pi}{4}$. In other words, the connected correlator only needs a very small time $\delta t \ll 1$ to grow from almost zero to a significant value. It gives evidence that creation of multipartite states can be expedited by small initial correlations. We remark that while the exact correlator $u_n(Z_1, \ldots, Z_n)$

is negligible at any time before $\frac{\pi}{4}$, there may exist other sets of observables for which n-point connected correlators are non-negligible.

6.2.4 Discussion

Although the relation between genuine multipartite entanglement and multipartite connected correlations is simple for pure states, whether one can infer any information about genuine multipartite entanglement of a mixed state from its multipartite correlations is still an open question.

In our model, only short-range interactions between two sites are present. An immediate question is how the Lieb-Robinson bounds generalize to other models with long-range interactions or interaction terms which involve more than two sites.

Current techniques to measure multipartite connected correlators require statistics of all measurement outcomes that factor into Eq. (6.14). Connected correlators up to tenth order have been measured using this approach [136]. However, such a method is infeasible for connected correlators of very high order, as the number of disconnected correlators that must be measured grows exponentially with n. It is therefore an open question whether there exist experimentally accessible observables, e.g. magnetization [92], which manifest multipartite connected correlators directly.

6.3 Simulation of Local Observables

In this section, we use the Lieb-Robinson bounds to improve the estimation of local observables in time-evolved states. Given an initial state $|\psi\rangle$ and a power-law Hamil-

tonian H, we consider the task of estimating the expectation value of the time-evolved observable $\langle A(t)\rangle := \langle\psi| U(t)^\dagger A U(t) |\psi\rangle$, where $U(t)$ is the unitary generated by H at time t, for a local operator A. The ability to perform this task for any arbitrary local observable is equivalent to the ability to compute local density matrices of the time-evolved state $U(t)|\psi\rangle$ or the ability to sample local observables in $U(t)|\psi\rangle$.

A typical approach to estimating $\langle A(t)\rangle$ is as follows. First the unitary evolution $U(t)$ on the entire system is decomposed into a more tractable sequence of elementary unitaries that are supported on a smaller number of sites to produce an approximation to the time-evolved state $|\psi(t)\rangle$. The expectation value $\langle A(t)\rangle$ is then computed by simulating measurements of A on this state. The number of elementary unitaries in the decomposition of $U(t)$ typically increases with both time t and the number of sites N in the system.

However, in the Heisenberg picture, the intuition from the Lieb-Robinson bounds suggests that the dynamics of $A(t) = U(t)^\dagger A U(t)$ is mostly confined inside some light cones and, therefore, it should be sufficient to simulate the unitary generated by the Hamiltonian inside the light cones alone. The following result provides such an approximation.

Proposition 3. *Let H be a 2-local power-law Hamiltonian (i.e. the sets X in Eq. (4.19) obey $|X| = 2$) of exponent $\alpha > 2d + 1$, and H_r be a Hamiltonian constructed from H by taking only interaction terms supported entirely on sites inside a ball of radius $r \geq 4\bar{v}t \geq 1$ around the support of the single-site operator A (where $\bar{v}$ is the same constant as in Proposition 2). Let $A(t)$ and $\tilde{A}(t)$ be the versions of A evolved for time t*

under H and H_r, respectively. Then there exists $0 < K < \infty$ such that

$$\|A(t) - \tilde{A}(t)\| = K \frac{t^{d+2} \log^{2d} r}{r^{\alpha - d}}. \tag{6.28}$$

Proof. Without loss of generality, assume that A is initially supported at the origin. Using the triangle inequality, we bound the difference between $A(t)$ and $\tilde{A}(t)$:

$$\|A(t) - \tilde{A}(t)\| \leq \int_0^t \mathrm{d}s \| \Big[H - H_r, \tilde{A}(s) \Big] \| \leq \int_0^t \mathrm{d}s \sum_{i:\mathrm{dist}(i,0) \leq r} \| \left[\sum_{j:\mathrm{dist}(j,0) > r} h_{i,j}, \tilde{A}(s) \right] \|. \tag{6.29}$$

We then use the bounds in Refs. [18, 137] to bound the commutator norm:

$$\| \left[\sum_{j:\mathrm{dist}(j,0) > r} h_{i,j}, \tilde{A}(s) \right] \|. \tag{6.30}$$

For that, we separate the sums over i into terms corresponding to i's inside and outside the linear light cone defined by $\mathrm{dist}(i,0) = 2\bar{v}s$.

For i such that $\mathrm{dist}(i,0) \leq 2\bar{v}s$, we simply use another triangle inequality for the sum over j and bound $\| \Big[h_{i,j}, \tilde{A}(s) \Big] \| \leq 2/\mathrm{dist}(i,j)^\alpha$. Note that in this case, we have $\mathrm{dist}(i,j) \geq \mathrm{dist}(j,0) - 2\bar{v}s \geq \mathrm{dist}(j,0)/2$. Therefore, we have

$$\sum_{j:\mathrm{dist}(j,0) > r} \sum_{i:\mathrm{dist}(i,0) \leq 2\bar{v}s} \| \Big[h_{i,j}, \tilde{A}(s) \Big] \| \leq 4^d 2^{\alpha+1} \bar{v}^d \sum_{j:\mathrm{dist}(j,0) > r} \frac{s^d}{\mathrm{dist}(0,j)^\alpha} \leq \frac{K s^d}{r^{\alpha - d}}, \tag{6.31}$$

for some constant $0 < K < \infty$. On the other hand, for i such that $r \geq \mathrm{dist}(i,0) > 2\bar{v}s$, we further divide into two cases: $s \geq 1$ and $s < 1$. For $s \geq 1$, we use Proposition 2 (note

that A is a single site operator):

$$\sum_{i:r\geq \text{dist}(i,0)>2\bar{v}s} \| \left[\sum_{j:\text{dist}(j,0)>r} h_{i,j}, \tilde{A}(s) \right] \| \leq K_1 \sum_{i:r\geq \text{dist}(i,0)>2\bar{v}s} \frac{1}{[r-\text{dist}(i,0)]^{\alpha-d}} \frac{s^{d+1}\log^{2d} r}{\text{dist}(i,0)^{\alpha-d}}$$
$$\leq K_2 s^{d+1} \frac{\log^{2d} r}{r^{\alpha-d}}, \tag{6.32}$$

where we have used Eq. (4.24) and defined another set of constants $0 < K_{1,2} < \infty$. Similarly, for $s < 1$, we use a bound from Ref. [18] to show that there exists $0 < K_3 < \infty$ such that

$$\sum_{j:\text{dist}(j,0)>r} \sum_{i:\text{dist}(i,0)>2\bar{v}s} \| \left[h_{i,j}, \tilde{A}(s) \right] \| \leq \frac{K_3}{r^{\alpha-d}}. \tag{6.33}$$

Substituting Eq. (6.31), Eq. (6.32), and Eq. (6.33) into Eq. (6.29) and integrating over time, we obtain Eq. (6.28). □

We now analyze the cost of estimating $\langle A(t) \rangle$ using quantum algorithms, although we note that Proposition 3 applies equally well to classical simulation algorithms. For simplicity, we assume that the Hamiltonian is time-independent in the following discussion. In order for the error of the approximation to be at most a constant, we choose

$$r \propto \max \left\{ t^{\frac{d+2}{\alpha-d}} \log t, t \right\}. \tag{6.34}$$

Therefore, to estimate $\langle A(t) \rangle$, it is sufficient to simulate the evolution of $\tilde{A}(t)$ on $N_r \propto r^d$ sites (instead of simulating the entire lattice).

We then compute $\left\langle \tilde{A}(t) \right\rangle$ by simulating $\mathrm{e}^{-iH_r t}$ using one of the existing quantum

algorithms. Using the pth-order product formula for simulating power-law Hamiltonians [77], we need

$$O\left((N_r t)^{\frac{\alpha}{\alpha-d}+o(1)}\right) = \max\left\{O\left(t^{\frac{\alpha(\alpha+d+d^2)}{(\alpha-d)^2}+o(1)}\right), O\left(t^{\frac{\alpha(1+d)}{\alpha-d}+o(1)}\right)\right\} \tag{6.35}$$

elementary quantum gates, where $o(1)$ denotes p-dependent constants that can be made arbitrarily small by increasing the order p. For all $\alpha > 2d+1$, this gate count is less than the estimate without using the Lieb-Robinson bound in Ref. [77]. In particular, in the limit $\alpha \to \infty$, the gate count reduces to $O\left(t^{1+d+o(1)}\right)$, which corresponds to the space-time volume inside a linear light cone.

We note that in estimating the gate count for computing $\langle A(t)\rangle$, we have implicitly assumed that we have access to many quantum copies of the initial state $|\psi\rangle$. However, in scenarios where only a classical description of $|\psi\rangle$ is provided, we need to add the cost of preparing $|\psi\rangle$ to the total gate count of the simulation.

6.4 Quantum State Transfer

In this section, we discuss the implication of the hierarchy of light cones for quantum state transfer. An immediate consequence of the Frobenius light cone is that the Lieb-Robinson light cone *is not relevant* for infinite temperature many-body quantum chaos and the growth of operators. A more practical application of the Frobenius light cone are tighter constraints on at least two different kinds of quantum state transfer. For simplicity, we assume that $\dim(\mathcal{H}_i) = 2$, and denote $|0_i\rangle$ and $|1_i\rangle$ as the eigenstates of the Pauli matrix Z_i on $\mathcal{H}_i$.

A *universal* notion of quantum state transfer from $i \in \Lambda$ to $j \in \Lambda$, which is independent of the background state, is to demand that there exists a Hamiltonian protocol $H(t)$ and a time $\tau \in \mathbb{R}$ such that

$$X_i^\alpha(\tau) = X_j^\alpha. \tag{6.36}$$

It is obvious that Theorem 4 constrains the time at which Eq. (6.36) may hold; hence Eq. (6.36) cannot be performed at a time τ which scales slower than linearly in the distance $\mathcal{D}(i,j)$ when $\alpha > \frac{5}{2}, d = 1$.

Another interesting scenario arises when we restrict to a time evolution operator $U(t)$ that obeys

$$U(t)\left(\bigotimes_{k\in\Lambda} |0_k\rangle\right) = \bigotimes_{k\in\Lambda} |0_k\rangle. \tag{6.37}$$

Many protocols, including our own (Theorem 6) and that of Ref. [24], are compatible with (6.37). With Eq. (6.37), we now consider a quantum system whose initial condition is

$$|\psi(0)\rangle := |\phi_i\rangle \otimes \bigotimes_{k\in\Lambda-\{i\}} |0_k\rangle. \tag{6.38}$$

for arbitrary $|\phi_i\rangle \in \mathcal{H}_i$. Our goal is to find a time evolution operator $U(t)$ and a time τ, such that $|\psi(t)\rangle = U(t)|\psi(0)\rangle$ and

$$\langle\psi(\tau)|Z_j|\psi(\tau)\rangle = \langle\phi_i|Z_i|\phi_i\rangle. \tag{6.39}$$

In particular, the probability of measuring a 0 or 1 on site j at time τ is given by the

probability of measuring it at time $t = 0$ on site i. This property must hold for *all* $|\phi_i\rangle$ for a fixed $U(t)$, since the protocol must be able to transfer arbitrary states.

Corollary 2. *Let* $\frac{3}{2} < \alpha \in \mathbb{R}$ *and* $x = \mathcal{D}(i,j)$. *Assuming (6.37), there exist* $0 < K, K' < \infty$ *such that any state transfer algorithm runs in a time* τ *obeying*

$$\tau > K \times \begin{cases} x & \alpha > \frac{5}{2} \\ x^{\alpha-3/2}(1 + K' \log x)^{-1} & \frac{3}{2} < \alpha \leq \frac{5}{2} \end{cases} . \tag{6.40}$$

Proof. We begin by observing that we may assume $|\phi_i\rangle = |1_i\rangle$ without loss of generality, since Eq. (6.39) is trivially obeyed by Eq. (6.37). Now the proof largely mirrors that of Theorem 4. Without loss of generality, we may define lattice sites such that $i = 0$ and $j > 0$, as explained above. Define

$$|S\rangle := \left(\bigotimes_{k \in S} |1_k\rangle \otimes \bigotimes_{k \in S^c} |0_k\rangle \right), \tag{6.41}$$

and the observable $\mathcal{F}$ which acts on the mutual eigenbasis of Z_i as

$$\mathcal{F}|S\rangle := \mathcal{F}_S|S\rangle, \tag{6.42}$$

for any $S \subseteq \Lambda$; here $\mathcal{F}_S$ is given by Eq. (4.66) when $\alpha > \frac{5}{2}$ and Eq. (4.82) when $\frac{3}{2} < \alpha \leq \frac{5}{2}$. For simplicity we only describe explicitly the case $\alpha > \frac{5}{2}$, as the other case follows

from identical considerations. We evaluate

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\mathcal{F}|\psi(t)\rangle\right| \leq |-\mathrm{i}\langle\psi(t)|[\mathcal{F}, H(t)]|\psi(t)\rangle| \leq \|[\mathcal{F}, H(t)]\|_\infty. \tag{6.43}$$

As before, our goal is to show that $\|[\mathcal{F}, H(t)]\|_\infty < \infty$. Since H is 2-local, we know that $H_{ij}(t)|0_i\rangle|0_j\rangle \propto |0_i\rangle|0_j\rangle$ by Eq. (6.37). This implies that, as before $[\mathcal{F}, H]$ can only be non-vanishing when H serves to either add a new $|1\rangle$ to the right end of the state, or delete the right-most $|1\rangle$. Hence $\langle S|[\mathcal{F}, H(t)]|Q\rangle \neq 0$ only if $|S - S \cap Q| \leq 1$ and $|Q - S \cap Q| \leq 1$. We define the matrix $M_{SQ} := \sup\langle S|[\mathcal{F}, H(t)]|Q\rangle$, which equals

$$M_{SQ} = M_{QS} := \begin{cases} A|\mathcal{F}_S - \mathcal{F}_Q|^{2-\alpha} & S = Q \cup \{m\} \text{ or } Q = S \cup \{m\} \\ A|\mathcal{F}_S - \mathcal{F}_Q|^{1-\alpha} & \text{there exists } R \text{ with } S = R \cup \{m\} \text{ and } Q = R \cup \{n\}, \\ & \text{and } Q \neq S \text{ and } \mathcal{F}_R < \min(\mathcal{F}_S, \mathcal{F}_Q), \\ 0 & \text{otherwise} \end{cases}. \tag{6.44}$$

We bound the maximal eigenvalue of M using the Collatz-Wieland inequality Eq. (4.74), using the trial vector φ_S given Eq. (4.76). Observe that the first line of Eq. (6.44) is identical to Eq. (4.71); as such these terms in $M_{SQ}\varphi_Q$ are bounded by a constant as before. The new terms we must deal with arise from the second line of Eq. (6.44). If S is given by Eq. (4.75), we find that

$$\frac{1}{\varphi_S}\sum_{Q:|Q|=|S|} M_{SQ}\varphi_Q < A\sum_{m=n_{\ell-1}+1}^{\infty}\left(\frac{n_\ell - n_{\ell-1}}{m - n_{\ell-1}}\right)^{\alpha-2}\frac{1-\delta_{m,n_\ell}}{(m-n_{\ell-1})^{\alpha-1}} < A_{\mathrm{st}}, \tag{6.45}$$

for some constant $0 < A_{\mathrm{st}} < \infty$, so long as $\alpha > \frac{5}{2}$. We conclude that M has a bounded

maximal eigenvalue, independently of the lattice size. We conclude there exists $0 < K < \infty$ such that $\langle\psi(t)|\mathcal{F}|\psi(t)\rangle \leq Kt$.

At time τ, we must have

$$|\psi(\tau)\rangle = |1_j\rangle \otimes |\psi'_{\Lambda-\{j\}}\rangle, \tag{6.46}$$

for arbitrary state $|\psi'\rangle$ acting on sites other than j. Therefore,

$$\langle\psi(\tau)|\mathcal{F}|\psi(\tau)\rangle \geq j. \tag{6.47}$$

Using Markov's inequality as in the proof of Theorem 4, we obtain Eq. (6.40). The case $\alpha < \frac{5}{2}$ is proved analogously. □

6.5 Simulation of a Free Particle

In this section, we discuss an immediate application of the free light cone (Theorem 5) in bounding the error made by approximating time evolution via a truncated, local Hamiltonian, analogous to the discussion of Section 6.3.

Corollary 3. *For any $i \in \Lambda$, define $B_i^r := \{j \in \Lambda : \mathcal{D}(j,i) \leq r\}$, and define $\tilde{H}(t)$ to be the restriction of a free bosonic Hamiltonian $H(t)$ [Eq. (4.95)] to $B_i^r \subset \Lambda$. Then for any $\varepsilon > 0$, there exists $0 < K, K' < \infty$ such that for times*

$$t < \frac{K'}{n} r^{\min(1,(\alpha-d-\varepsilon)/3)}, \tag{6.48}$$

we have

$$\|b_i^\dagger(t) - \tilde{b}_i^\dagger(t)\| \leq Kn^{3/2}\left(\frac{t}{r^{\alpha-d}} + \frac{t^{3/2}}{r^{(\alpha-d-\varepsilon)/2}}\right), \tag{6.49}$$

where the norm is estimated in the subspace that has at most $n \geq 1$ excitations across the lattice and $\tilde{b}_i^\dagger(t)$ denotes time evolution with the restricted Hamiltonian $\tilde{H}(t)$.

Proof. Without loss of generality, we assume $i = 0$, the origin. Observe that

$$\|b_0^\dagger(t) - \tilde{b}_0^\dagger(t)\| \leq \int_0^t \mathrm{d}s \|\left[\tilde{b}_0^\dagger(t), H(t) - \tilde{H}(t)\right]\| \tag{6.50}$$

$$\leq \int_0^t \mathrm{d}s \|\left[\tilde{b}_0^\dagger(t), \sum_{i:\mathrm{dist}(i,0)\leq r} \sum_{j:\mathrm{dist}(j,0)>r} h_{ij} b_i b_j^\dagger\right]\|. \tag{6.51}$$

Using Theorem 5,

$$\tilde{b}_0^\dagger(t) = \sum_{i:\mathrm{dist}(i,0)\leq r} f_i(t) b_i^\dagger, \tag{6.52}$$

where the coefficients $f_i(t)$ satisfy, for some $0 < C < \infty$ and arbitrary $\varepsilon > 0$,

$$\sum_{i:\mathrm{dist}(i,0)\geq x} |f_i(t)|^2 \leq \frac{Ct}{x^{\alpha-d-\varepsilon}}, \tag{6.53}$$

for all $x > 0$ and all t obeying Eq. (6.48).

We separate the sum over i in Eq. (6.51) according to $\mathrm{dist}(i,0) \leq r/2$ and $r/2 <$

$\mathrm{dist}(i,0) \leq r$. In the former case, we have

$$\begin{aligned}
\| \left[\tilde{b}_0^\dagger(t), \sum_{i:\mathrm{dist}(i,0)\leq r/2} \sum_{j:\mathrm{dist}(j,0)>r} h_{ij} b_i b_j^\dagger \right] \| &\leq 2\sqrt{n} \| \sum_{i:\mathrm{dist}(i,0)\leq r/2} \sum_{j:\mathrm{dist}(j,0)>r} h_{ij} b_i b_j^\dagger \| \\
&\leq 2n^{3/2} \max_{i:\mathrm{dist}(i,0)\leq r/2} \sum_{j:\mathrm{dist}(j,0)>r} \frac{1}{\mathrm{dist}(i,j)^\alpha} \leq \frac{C_1 n^{3/2}}{r^{\alpha-d}},
\end{aligned} \tag{6.54}$$

where $0 < C_1 < \infty$ is a constant. We have used the fact that $|h_{ij}| \leq 1/\mathrm{dist}(i,j)^\alpha$ and that $\mathrm{dist}(j,i) \geq r/2$ for all i such that $\mathrm{dist}(i,0) \leq r/2$.

On the other hand, for $r/2 < \mathrm{dist}(i,0) \leq r$,

$$\begin{aligned}
&\| \left[\tilde{b}_0^\dagger(t), \sum_{i:r/2\leq \mathrm{dist}(i,0)\leq r} \sum_{j:\mathrm{dist}(j,0)>r} h_{ij} b_i b_j^\dagger \right] \| \\
&\leq 2 \| \sum_{i:r/2<\mathrm{dist}(i,0)\leq r} f_i(t) b_i^\dagger \| \| \sum_{i:r/2<\mathrm{dist}(i,0)\leq r} \sum_{j:\mathrm{dist}(j,0)>r} h_{ij} b_i b_j^\dagger \| \\
&\lesssim \sqrt{\sum_{i:r/2<\mathrm{dist}(i,0)\leq r} |f_i(t)|^2 n} \left(\max_{i:r/2<\mathrm{dist}(i,0)\leq r} \sum_{j:\mathrm{dist}(j,0)>r} \frac{n}{\mathrm{dist}(i,j)^\alpha} \right) \\
&\leq C_2 \frac{n^{3/2} t^{1/2}}{r^{(\alpha-d-\varepsilon)/2}},
\end{aligned} \tag{6.55}$$

for $0 < C_2 < \infty$. Replacing Eqs. (6.54) and (6.55) into Eq. (6.51) and integrating over time, we arrive at Eq. (6.49). □

6.6 Classical Complexity of Boson Sampling

In this section, we discuss the implications of the free light cone and the corresponding fast free-particle state transfer protocol on the classical complexity of boson sampling.

The boson sampling problem was proposed by Aaronson and Arkhipov [82] as a potential candidate for the demonstration of quantum supremacy. While simulating the dynamics of bosons hopping on a lattice is generally a difficult task for classical computers, early-time evolutions where the bosons do not have enough time to hop too far from their initial positions can be simulated efficiently [16, 83, 138]. In particular, Ref. [16] considered a scenario where bosons were initially located at equal distances on a lattice and allowed to move in the lattice using only nearest-neighbor hoppings. Using the Lieb-Robinson bounds, the authors constructed an early-time classical sampler that efficiently captures the dynamics of the bosons up to time t_* that scales polynomially with the system size and thereby demonstrated a dynamical phase transition in the computational complexity.

The early-time classical sampler was later generalized to more complicated systems with power-law hoppings [83]. However, the easiness timescale t_* in this case only scales polynomially with the system size for $\alpha > 2d$ and scales logarithmically with the system size when $d + 1 < \alpha < 2d$. In this section, we show that the tight free-particle bound in this chapter immediately imply that t_* scales polynomially with the system size for all $\alpha > d$, i.e. an exponentially longer easiness timescale in the regime $\alpha \in (d, 2d]$ compared to the previous results [83].

For pedagogical reasons, we only describe here the high-level ideas behind the construction of the early-time boson sampler and argue for its efficiency using the technical results of Ref. [83]. We consider N bosons hopping on a d-dimensional lattice under the

Hamiltonian

$$H(t) = \sum_{i,j} J_{i,j}(t) b_i^\dagger b_j, \tag{6.56}$$

where b_i is the bosonic annihilation operator on site i, $J_{i,j}(t) \leq 1/\text{dist}(i,j)^\alpha$ are the hopping strengths, and the sums are over all sites i, j on the lattice. We assume that the lattice has $M \propto N^\beta$ sites in total, where $\beta \geq 1$ is a constant.

The bosons are initially located on evenly spaced sites on the lattice so that the minimum distance between nearest occupied sites is $2L \propto (M/N)^{1/d} \propto N^{(\beta-1)/d}$, as shown in Figure 6.5. Denote these initial positions by $j_1, \ldots, j_N$. We can write the initial state in terms of the creation operators:

$$|\psi(0)\rangle = \prod_{k=1}^{N} b_{j_k}^\dagger |\text{vac}\rangle , \tag{6.57}$$

where $|\text{vac}\rangle$ is the vacuum state.

The aim of boson sampling is to sample the positions of the bosons at a later time t. The idea of the early-time boson sampler in Refs. [16, 83] is that each boson primarily hops within its causal light cone, i.e. a bubble of radius $r(t)$ centered on its initial position. For small enough time, $r(t) < L$ and the bosons do not interfere with each other. The state of the system at this time can be approximated by a product state over the bubbles and therefore the positions of the bosons can be efficiently simulated by simulating the dynamics of each boson independently.

Let $U(t) = \mathcal{T} \exp(-i \int_0^t \mathrm{d}s H(s))$ be the evolution unitary generated by H at time

t. By inserting pairs of $\mathbb{I} = U^\dagger U$ in between the creation operators, the state of the system at time t can be written as

$$|\psi(t)\rangle = \prod_{k=1}^{N} U(t) b_{j_k}^\dagger U^\dagger(t) \, |\text{vac}\rangle \, . \tag{6.58}$$

Here, the evolution of the state can be simplified into independent evolutions of the creation operators $b_{j_k}^\dagger(t) \equiv U(t) b_{j_k}^\dagger U^\dagger(t)$. Using our free-particle bound in Theorem 5, we can approximate $b_{j_k}^\dagger(t)$ by its evolution within a light cone originated from j_k:

$$b_{j_k}^\dagger(t) \approx U_k(t) b_{j_k}^\dagger U_k^\dagger(t) \equiv \tilde{b}_{j_k}^\dagger(t), \tag{6.59}$$

where $U_k(t) = \mathcal{T} \exp[-i \int_0^t \mathrm{d}s H_k(s)]$ and H_k is the Hamiltonian constructed from H by taking only the hoppings between sites that are at most a distance L from j_k. Using Corollary 3, the error of this approximation is $\mathcal{O}\left((Nt)^{3/2}/L^{(\alpha-d-\varepsilon)/2}\right)$, where ε is an arbitrarily small positive constant and we have assumed $t \geq 1$ without loss of generality. Repeating the approximations for all $k = 1, \ldots, N$, we thereby show that the state $|\psi(t)\rangle$ is approximately $|\phi(t)\rangle = \prod_k \tilde{b}_{j_k}(t) \, |\text{vac}\rangle$.

Since the operators $\tilde{b}_{j_k}(t)$ are supported on distinct regions, the bosons from different regions do not interfere with each other. Therefore the probability distribution for the positions of the bosons in $|\phi(t)\rangle$ is simply the product of probability distributions of each boson hopping independently. Thus, at later time, the positions of the bosons in $|\phi(t)\rangle$ can be efficiently sampled on a classical computer.

Note that the state $|\phi(t)\rangle$ only approximates $|\psi(t)\rangle$ up to some time t_*. To estimate

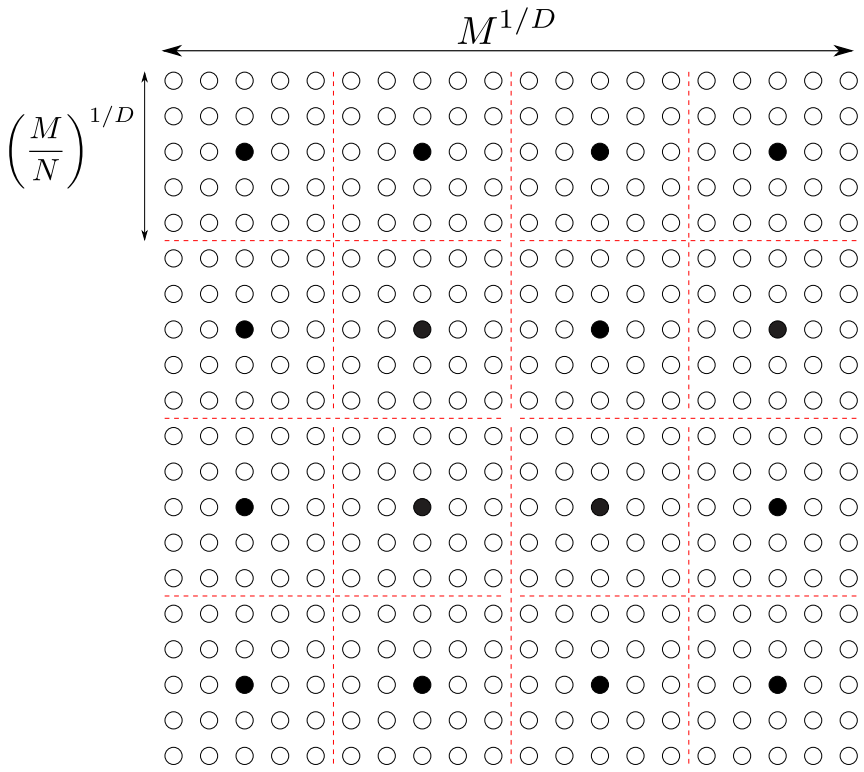

Figure 6.5: A depiction of the initial state in Ref. [16]. Empty circles represent empty lattice site and filled circles represent initially occupied sites.

t_*, we calculate the total error of the approximation. A simple calculation [83] shows that the total error of approximating the N original bosons $\{b^\dagger(t)\}$ by the confined ones $\{\tilde{b}^\dagger(t)\}$ would be $\mathcal{O}\left(N^{5/2}t^{3/2}/L^{(\alpha-d-\varepsilon)/2}\right)$—$N$ times the error of approximating each $b^\dagger(t)$ by the corresponding $\tilde{b}^\dagger(t)$.

Requiring that the total error of the approximation is at most a small constant, we obtain

$$t_* \propto L^{\frac{\alpha-d-\varepsilon}{3}} N^{-\frac{5}{3}} \propto N^{\frac{(\beta-1)(\alpha-d-\varepsilon)}{3d}-\frac{5}{3}}, \tag{6.60}$$

where we have replaced $L \propto N^{(\beta-1)/d}$ from our assumption. Therefore, by choosing a small enough ε, the easiness time t_* increases polynomially with N for all $\alpha > d(1+\frac{5}{\beta-1})$.

In particular, the condition becomes $\alpha > d$ in the limit of large β. Therefore, our free-particle bound has improved the easiness time t_* exponentially compared to Ref. [83] in the regime $\alpha \in (d, 2d]$.

6.7 Generation of Topologically Ordered States

In this section, we study the minimum time it takes to create topologically ordered states from topologically trivial ones. Before we present our result, we shall define topologically ordered states and topologically trivial states following the definitions in Refs. [36, 139]. Suppose that the finite lattice Λ has diameter L and consists of $O(L^d)$ sites. Let $\{|\psi_1\rangle, \cdots |\psi_k\rangle\}$ be a set of orthonormal quantum many-body states and define

$$\varepsilon = \sup_{\mathcal{O}} \max_{1 \leq i,j \leq k} \left\{ \left| \langle \psi_i | \mathcal{O} | \psi_i \rangle - \langle \psi_j | \mathcal{O} | \psi_j \rangle \right|, 2 \left| \langle \psi_i | \mathcal{O} | \psi_j \rangle \right| \right\}, \tag{6.61}$$

where the supremum is taken over unit-norm operators $\mathcal{O}$ supported on a subset of the lattice with diameter $l \leq bL$ for a constant b. Roughly speaking, ε quantifies the ability to distinguish between the states $\{|\psi_1\rangle, \ldots, |\psi_k\rangle\}$ using observables that are supported on only a fraction of the lattice. We say that the states are topological with diameter bL if there exist constants $b, c, \beta > 0$ such that $\varepsilon \leq cL^{-\beta}$, and are trivial if ε is independent of L for all b [140]. We now use the Lieb-Robinson bound to bound the minimum time it takes to convert between topological and trivial states.

Proposition 4. *Consider a time-dependent Hamiltonian H with long-range interactions of exponent α on Λ. Let $U(t)$ be the evolution unitary generated by H at time t, let*

$\{|\psi_1\rangle, \cdots |\psi_k\rangle\}$ be a set of topologically ordered states, and let $\{|\phi_1\rangle, \cdots |\phi_k\rangle\}$ be a set of topologically trivial states. If $\alpha > 2d+1$ and there is a time $0 < \tau < \infty$ such that $|\psi_i\rangle = U(\tau)|\phi_i\rangle$, then there exists an L-independent constant $0 < K < \infty$ such that $\tau > K\tau^*$, where

$$\tau^* := \begin{cases} L & \text{when } \alpha > 3d+1, \\ L^{\frac{\alpha-2d}{d+1}} / \log^{2d} L & \text{when } 2d+1 \leq \alpha \leq 3d+1. \end{cases} \tag{6.62}$$

Proof. Suppose $\{|\psi_k\rangle\}$ are topologically ordered with diameter $l' = bL$. Consider an arbitrary operator $\mathcal{O}$ with a support diameter at most $l = l'/2$ and let $\mathcal{O}(t) \equiv U(t)\mathcal{O}U^\dagger(t)$ be the evolved version of $\mathcal{O}$. We further introduce $\mathcal{O}(t, l') = \mathrm{tr}_{\mathcal{B}_{l'}^c}\mathcal{O}(t)$ as the version of $\mathcal{O}(t)$ truncated to a ball $\mathcal{B}_{l'}$ of diameter l'. Using the triangle inequality, we have

$$|\langle\phi_i|\mathcal{O}|\phi_i\rangle - \langle\phi_j|\mathcal{O}|\phi_j\rangle| = |\langle\psi_i|\mathcal{O}(\tau)|\psi_i\rangle - \langle\psi_j|\mathcal{O}(\tau)|\psi_j\rangle| \tag{6.63}$$

$$\leq 2\left\|\mathcal{O}(\tau) - \mathcal{O}(\tau, l')\right\| + |\langle\psi_i|\mathcal{O}(\tau, l')|\psi_i\rangle - \langle\psi_j|\mathcal{O}(\tau, l')|\psi_j\rangle|\,. \tag{6.64}$$

By our assumptions on the topological order of $|\psi_k\rangle$ and absence of topological order in $|\phi_k\rangle$, there exist an operator $\mathcal{O}$ and constants $0 < \beta, a_{1,2} < \infty$ such that

$$a_1 - \frac{a_2}{L^\beta} \leq 2\left\|\mathcal{O}(\tau) - \mathcal{O}(\tau, l')\right\|. \tag{6.65}$$

On the other hand, using Proposition 2 for $\alpha > 2d+1$ and $\tau < L/\bar{v}$, where $\bar{v}$ is a

constant, there exists $0 < C_{1,2} < \infty$ such that

$$\|\mathcal{O}(t) - \mathcal{O}(\tau, l')\| \leq C_1 L^d \frac{\tau^{d+1} \log^{2d} l'}{l'^{\alpha-d}} = C_2 \tau^{d+1} \frac{\log^{2d} L}{L^{\alpha-2d}}, \tag{6.66}$$

where the factor L^d accounts for the support size of $\mathcal{O}$. For all $\alpha > 2d+1$, Eq. (6.66) vanishes as L increases, in contradiction with Eq. (6.65), unless $\tau = O(\tau^*)$. This completes the proof. □

6.8 Clustering of Correlations

In addition to the dynamics of quantum systems, the Lieb-Robinson bounds also have implications for the eigenstates of a Hamiltonian. In Ref. [18], the authors show that if a time-independent power-law Hamiltonian with an exponent α has spectral gap $\Delta > 0$, the correlations between distant sites in the ground state of the system also decay with the distance as a power law with an exponent lower bounded by

$$\alpha' = \frac{\alpha}{1 + \tilde{v}\Delta^{-2}}, \tag{6.67}$$

where $\tilde{v}$ is a constant that depends on α.

The bound in Ref. [18] has a undesirable feature: for a given value of α, varying the gap Δ also changes the minimum exponent α'. Although this leads to an intuitive implication that larger energy gaps result in faster correlation decay, there is no known example where ground state correlations decay at a slower rate than a power law with an exponent α. Indeed, we shall show that the cause of this problem is tied to the previous

lack of an algebraic light cone in the quench dynamics. In particular, by using the Lieb-Robinson bounds with algebraic light cones [19, 22, 23, 74, 141], we show for all $\alpha > 2d$ that the ground state correlations must decay as a power law with the exponent lower bounded by the exponent of the Hamiltonian.

Proposition 5. *Let H be a power-law Hamiltonian with an exponent α; let A, B be local operators obeying $\|A\|, \|B\| \leq 1$, supported on X, Y such that $|X| = |Y| = 1$ and dist$(X, Y) = r > 0$. Assume that H has a unique ground state $|\psi_0\rangle$ and spectral gap Δ to the first excited state. Define $\mathcal{C}(r) \coloneqq \langle\psi_0|AB|\psi_0\rangle - \langle\psi_0|A|\psi_0\rangle\langle\psi_0|B|\psi_0\rangle$ to be the connected correlator between A, B in the ground state. Then*

$$|\mathcal{C}(r)| \leq \left[\frac{2^{\gamma-1}c\Gamma(\frac{\gamma}{2})}{\pi}\frac{\alpha^{\gamma/2}}{\Delta^{\gamma}} + 1\right]\frac{\log^{\gamma/2} r}{r^{\alpha}}, \tag{6.68}$$

where c is a constant independent of α, $\gamma = \alpha(\alpha - d + 1)/(\alpha - 2d)$, and $\Gamma(\cdot)$ is the Gamma function.

Proof. First we rewrite

$$\mathcal{C}(r) = \sum_{k>0} \langle\psi_0| A |\psi_k\rangle \langle\psi_k| B |\psi_0\rangle , \tag{6.69}$$

where the sum is over the excited states $|\psi_k\rangle$ of the Hamiltonian. Our strategy is to relate $\mathcal{C}(r)$ to the commutator norm $\|[A(t), B]\|$, which we then bound using a Lieb-Robinson bound. To relate $\mathcal{C}(r)$ to $\|[A(t), B]\|$, it is natural to first consider the value of $[A(t), B]$

in the ground state, whose magnitude is bounded by $\|[A(t), B]\|$:

$$\langle\psi_0|[A(t), B]|\psi_0\rangle = \langle\psi_0|A(t)B|\psi_0\rangle - \text{h.c.} = \sum_{k>0} \mathrm{e}^{iE_k t}\langle\psi_0|A|\psi_k\rangle\langle\psi_k|B|\psi_0\rangle - \text{h.c.}, \tag{6.70}$$

where E_k are the eigenvalues of the Hamiltonian and we have set ground state energy $E_0 = 0$ so that $E_k > 0$ for all $k > 0$. Note that the $k = 0$ terms cancel between the first term and its Hermitian conjugate.

By observation, we note that if we could replace the terms $\mathrm{e}^{iE_k t}$ in Eq. (6.70) by a unit step function $\Theta(E_k)$ that satisfies $\Theta(E_k) = 1$ and $\Theta(-E_k) = 0$, we immediately obtain the expression of $\mathcal{C}(r)$ in Eq. (6.69). In fact, this replacement is easy to achieve using the identity

$$\lim_{\varepsilon\to 0^+} \frac{1}{2\pi i}\int_{-\infty}^{\infty} \mathrm{d}t \frac{\mathrm{e}^{iE_k t}}{t - i\varepsilon} = \Theta(E_k). \tag{6.71}$$

Therefore, we have

$$\lim_{\varepsilon\to 0^+} \frac{1}{2\pi i}\int_{-\infty}^{\infty} \mathrm{d}t \frac{\langle\psi_0|[A(t), B]|\psi_0\rangle}{t - i\varepsilon} = \mathcal{C}(r), \tag{6.72}$$

and we obtain the relation

$$|\mathcal{C}(r)| = \left|\lim_{\varepsilon\to 0^+} \frac{1}{2\pi i}\int_{-\infty}^{\infty} \mathrm{d}t \frac{\langle\psi_0|[A(t), B]|\psi_0\rangle}{t - i\varepsilon}\right| \le \frac{1}{\pi}\int_0^{\infty} \mathrm{d}t \frac{\|[A(t), B]\|}{t}. \tag{6.73}$$

Unfortunately, this relation is not useful; the right-hand side of Eq. (6.73) diverges even when the commutator $\|[A(t), B]\|$ does not increase with time. The failure of such

a simple treatment is not surprising as we have not used the crucial assumption on the existence of a finite energy gap ($E_k \geq \Delta$).

Intuitively, to make the integral in Eq. (6.73) converge, we can multiply the integrand by a function that decays quickly with t. A natural choice is a Gaussian function $\mathrm{e}^{-(\upsilon t/2)^2}$, where $\upsilon > 0$ is an adjustable parameter; it decays with time quickly enough to make the integral converge and its Fourier transformation is rather easy to handle. By multiplying this function to the integrand in Eq. (6.71), we arrive at a convolution of the step function with the Gaussian function:

$$\lim_{\varepsilon\to 0^+} \frac{1}{2\pi i}\int_{-\infty}^{\infty} \mathrm{d}t \frac{\mathrm{e}^{iE_k t}\mathrm{e}^{-(\upsilon t/2)^2}}{t-i\varepsilon} = \frac{1}{\sqrt{\pi}\upsilon}\int_{-\infty}^{\infty} \Theta(E_k - E)\mathrm{e}^{-E^2/\upsilon^2} dE =: f(E_k). \tag{6.74}$$

It is easy to verify that $f(E_k) = 1 - g(E_k)$ and $f(-E_k) = 0 + g(E_k)$ for some positive function $g(E_k) \leq \frac{1}{2}\mathrm{e}^{-(E_k/\upsilon)^2}$. Thus, $f(E_k)$ closely resembles the step function $\Theta(E_k)$, albeit with a smoother transition from 0 to 1.

Inserting this convolution into Eq. (6.72), we have:

$$\lim_{\varepsilon\to 0^+} \frac{1}{2\pi i}\int_{-\infty}^{\infty} \mathrm{d}t \frac{\langle\psi_0|[A(t),B]|\psi_0\rangle \mathrm{e}^{-(\upsilon t/2)^2}}{t-i\varepsilon} = \mathcal{C}(r) - \sum_{k>0} g(E_k)[\langle\psi_0|A|\psi_n\rangle\langle\psi_n|B|\psi_0\rangle + \text{h.c.}]. \tag{6.75}$$

Using a Cauchy-Schwarz inequality, we can then bound the absolute value of the sum over k in the right-hand side by

$$\sum_{k>0} 2g(E_k)\,|\langle\psi_0|A|\psi_n\rangle\langle\psi_n|B|\psi_0\rangle| \leq \mathrm{e}^{-(\Delta/\upsilon)^2}, \tag{6.76}$$

where we have used that $E_k \geq \Delta$. Thus we arrive at our desired relation:

$$|\mathcal{C}(r)| \leq \frac{1}{\pi} \int_0^\infty \mathrm{d}t \frac{\mathrm{e}^{-(vt/2)^2}}{t} \|[A(t), B]\| + \mathrm{e}^{-(\Delta/v)^2}. \tag{6.77}$$

Finally, we bound the commutator norm using the Lieb-Robinson bound in Ref. [19],

$$\|[A(t), B]\| \leq c \frac{t^\gamma}{r^\alpha}, \tag{6.78}$$

where c is a constant and $\gamma = \alpha(\alpha - d + 1)/(\alpha - 2d)$. We obtain:

$$|\mathcal{C}(r)| \leq \frac{2^{\gamma-1} c \Gamma(\frac{\gamma}{2})}{\pi} \frac{1}{v^\gamma r^\alpha} + \mathrm{e}^{-(\Delta/v)^2}, \tag{6.79}$$

where $\Gamma(\cdot)$ is the Gamma function. By choosing $v = \Delta/\sqrt{\alpha \log r}$, we get

$$|\mathcal{C}(r)| \leq \left[\frac{2^{\gamma-1} c \Gamma(\frac{\gamma}{2})}{\pi} \frac{\alpha^{\gamma/2}}{\Delta^\gamma} + 1 \right] \frac{\log^{\gamma/2} r}{r^\alpha}. \tag{6.80}$$

Therefore, the correlators in the ground state of a power-law Hamiltonian with $\alpha > 2d$ also decay with the distance as a power law (up to a logarithmic correction) with the same exponent α as that of the Hamiltonian. In particular, this exponent is independent of the energy gap Δ, in contrast to the previous result in Ref. [18]. □

Note that in Eq. (6.79), we have used an algebraic light cone bound from [19] instead of the tighter bounds in recent works [22, 23, 74, 141], because the bounds in Refs. [22, 74, 141] decay with the distance slower than $1/r^\alpha$ while the bound in Ref. [23] does not hold for $2d < \alpha \leq 2d + 1$.

6.9 Heating Rate in Periodically Driven Systems

Periodically driven systems can host interesting non-equilibrium physics, such as Floquet topological insulators [142], time crystals [143, 144], and anomalous Floquet phases [145]. However, most driven systems eventually heat up to equilibrium, infinite-temperature states, erasing the interesting features in the process.

The timescale before heating becomes appreciable in periodically driven systems is known as the *heating time*, and it generally exhibits a nontrivial dependence on the frequency of the drive, ω. Previous works [137, 146, 147, 148] established that finite-range interacting systems under rapid, local [2], periodic drives could not heat up until after a time $t_* = e^{\mathcal{O}(\omega)}$ that is exponentially long in the drive frequency ω. This slow *heating rate* stems at least in part from the locality of the interactions, which constrains the probability that distant particles collectively absorb an energy quantum $\hbar\omega$.

This result also applies to systems with long-range interactions that decay with the distance r, e.g. as a power-law $1/r^\alpha$. Such systems are of great interest as they can be implemented in a wide variety of experiments, such as trapped ions [49, 50], Rydberg atoms [55], ultracold atoms and molecules [53, 59], nitrogen-vacancy centers [58], and superconducting circuits [114]. On the theoretical side, for spin systems with disordered, sign-changing power-law couplings, Ref. [149] demonstrated the exponentially-suppressed heating rate when $\alpha > d/2$, where d is the dimensionality of the system. Furthermore, Ref. [137] proved an exponential heating time $t_* = e^{\mathcal{O}(\omega)}$ for general power-law interactions with $\alpha > 2d$. In contrast, for $d < \alpha < 2d$, Ref. [137] only obtained a linear

[2] Refs. [137, 146, 147, 148] assumed that the drive is a sum of local terms.

heating time $t_* = \mathcal{O}(\omega)$, while numerical evidence [150] suggests that the heating time is still exponential within this regime of α.

In this section, we study the heating time in periodically driven, power-law interacting systems with $\alpha > d$ from two different perspectives. Within linear response theory, we show that such systems only heat up after some time exponentially large as a function of the drive frequency. This result mirrors the statement established for finite-range interactions in Ref. [146] and extends Ref. [149] to systems without disorder (though at the expense of a smaller range of valid α). The result also matches the numerical evidence in Ref. [150]. For generic periodically driven, power-law interacting systems that may not obey the linear response theory—such as those under a strong drive—we generalize Ref. [147] and construct an effective time-independent Hamiltonian with power-law interactions. This Hamiltonian closely describes the dynamics of the driven system up to time t_*, where t_* is exponentially large as a function of the drive frequency. We thereby show that the system cannot heat up until at least after this timescale.

We note that, although our generalization of Ref. [147] is different from Ref. [137], it is similar in spirit to their arguments. While Ref. [137] mainly focused on finite-range interactions, their construction of the effective Hamiltonian by truncating the Magnus series would also apply to power-law systems. However, our approach here also provides insights into the structure of the effective Hamiltonian. In particular, we show that the effective Hamiltonian is also power-law with the same exponent α as the driven Hamiltonian. Furthermore, we prove a stronger, albeit still exponential in ω, bound on the heating time than one would get from the argument in Ref. [137]. This improvement relies on the use of state-of-the-art Lieb-Robinson bounds [21, 141], which we develop for this pur-

pose. In particular, through a new technique, we generalize the bound in Ref. [21] from two-body to many-body interactions.

Similarly to Ref. [137], our construction requires $\alpha > 2d$, in contrast to the numerical evidence in Ref. [150] and to the wider range of validity $\alpha > d$ found in the linear response theory. Because both Ref. [137] and this section crucially rely on Lieb-Robinson bounds to prove that the heating time is at least exponential in ω, we conjecture that the aforementioned gap stems from the lack of a tight Lieb-Robinson bound for $\alpha > d$, and we show the gap would vanish if such a tight bound were to exist. While the linear response theory also utilizes Lieb-Robinson bounds, it has weaker assumptions and, therefore, does not require a tighter bound to achieve the desired result of exponentially suppressed heating for $\alpha > d$.

The remainder of the section is organized as follows. In Sec. 6.9.1, we provide definitions and describe the systems of interest. We then review various Lieb-Robinson bounds for power-law systems and extend two of them—including one with the tightest light cone known to date—to k-body interactions. In Sec. 6.9.2, we prove that in the linear response regime the heating time is at least exponential in ω for all $\alpha > d$. In Sec. 6.9.3, we provide a more general analysis using the Magnus-like expansion and existing Lieb-Robinson bounds to prove exponentially-long heating times for $\alpha > 2d$. We also conjecture a tight Lieb-Robinson bound that would extend this range of validity to $\alpha > d$. Finally, we summarize and discuss potential improvements in Sec. 6.9.4.

6.9.1 Setup and Definitions

We consider a system of N spins in a d-dimensional square lattice [3]. The system evolves under a periodic, time-dependent Hamiltonian $H(t)$ with period T, i.e. $H(t + T) = H(t)$. While the following analysis works for any $H(t)$ that is a sum of finite-body interactions, we assume that $H(t)$ is two-body for simplicity. Without loss of generality, we can write $H(t) = H_0 + V(t)$ as the sum of a time-independent part H_0 and a time-dependent part $V(t)$ such that $\frac{1}{T}\int_0^T V(t) = 0$. We further assume that H_0 and $V(t)$ are both power-law Hamiltonians with an exponent α.

Definition 1. *A Hamiltonian H on a lattice Λ is power-law with an exponent α and a local energy scale η if we can write $H = \sum_X h_X$, where h_X are Hamiltonians supported on subsets $X \subset \Lambda$, such that for any two distinct sites $i, j \in \Lambda$:*

$$\sum_{X \ni i,j} \|h_X\| \leq \frac{\eta}{dist(i,j)^\alpha}, \tag{6.81}$$

and the norm $\|h_{\{i\}}\| \leq \eta$ for all $i \in \Lambda$, where $\|\cdot\|$ denotes the operator norm and dist(i, j) the distance between sites i, j. In addition, we call $\sup_X |X|$ the local support size, where $|X|$ is the number of sites in X, and define $\|H\|_l = \sup_i \sum_{X \ni i} \|h_X\|$ to be the local norm of H.

In the following discussion, we assume $\eta = 1$, which sets the timescale for the dynamics of the system. In addition, we will occasionally write H instead of $H(t)$ for brevity.

[3] While our results are derived considering a simple square lattice, we believe that it is not difficult to extend them to other regular lattices.

Before discussing the linear response theory and the Magnus-like expansion, it is helpful to review the applicable Lieb-Robinson (LR) bounds for power-law interactions. We will also generalize several of them from two-body to arbitrary k-body interactions for all $k \geq 2$. In particular, we discuss the relations between the bounds in Refs. [18, 29], which imply logarithmic light cones for all $\alpha > d$, and the bounds in Refs. [19, 21, 141], which imply algebraic light cones for $\alpha > 2d$.

6.9.1.1 Lieb-Robinson Bounds for $\alpha > d$

First, we discuss the bounds in Refs. [18, 29], which are valid for all $\alpha > d$. Recall that LR bounds are upper bounds on the norm of the commutator $[A(t), B]$, where A, B are two operators supported on some subsets X, Y of the lattice, and $A(t)$ is the time-evolved version of A under a time-dependent Hamiltonian H. The minimum distance between a site in X and a site in Y is $r = \text{dist}(X, Y) > 0$. Since the sets X, Y are disjoint, $[A(0), B] = 0$ initially. As time grows, the operator $A(t)$ may spread to Y, making the commutator nontrivial.

The first LR bound for power-law interactions was proven in Ref. [18] by Hastings and Koma (HK):

$$\mathcal{C}(t, r) \equiv \|[A(t), B]\| \leq C\|A\|\|B\||X||Y|\frac{e^{vt}}{r^{\alpha}}, \tag{6.82}$$

where $r = \text{dist}(X, Y)$, $v \propto \eta$ is a constant that may depend on α, and C is a constant independent of the system. We shall also use the same C to denote different inconsequential prefactors. Setting the commutator norm to a constant yields the light cone $t \gtrsim \log r$,

which means it takes time at least proportional to $\log r$ for the commutator to reach a given constant value.

Technically, we can already use the HK bound in our later analysis of the heating time because it applies to k-body interactions for all k. However, this bound is loose for large α for two reasons: i) the velocity $v \propto 2^\alpha$ diverges for $\alpha \to \infty$, ii) the light cone is logarithmic for all α, which is unphysical since larger values of α correspond to shorter-range interactions and, therefore, we expect slower spreading of correlations. In particular, we expect the light cone to become linear for large enough α, given that the interactions are finite-range at $\alpha = \infty$.

Gong et al. [29] resolved the first challenge and derived a bound for two-body interactions:

$$\mathcal{C}(t,r) \leq C\|A\|\|B\||X||Y|\left(\frac{e^{vt}}{[(1-\mu)r]^\alpha} + e^{vt-\mu r}\right), \tag{6.83}$$

where $\mu \in (0,1)$ is an arbitrary constant. The light cone implied by this bound is still logarithmic, but the velocity v is finite for all α. Although the bound in Ref. [29] was derived only for two-body interactions, their proof applies to arbitrary k-body interactions, where k is a finite integer [See Appendix D.1 for a proof].

6.9.1.2 Lieb-Robinson Bounds for $\alpha > 2d$

In this section, we discuss the LR bounds for power-law interactions with $\alpha > 2d$. While the bounds in Ref. [18, 29] work for $\alpha > d$, they all have logarithmic light cones. For $\alpha > 2d$, it is possible to derive tighter algebraic light cones. The first such

bound was proven by Foss-Feig et al. [19] for two-body interactions (and generalized by Refs. [141, 151] to k-body interactions for all $k \geq 2$). A recent bound by Tran et al. [21], however, gives a tighter algebraic light cone. Here, we provide the generalization of that bound to k-body interactions:

$$\begin{aligned}\mathcal{C}(t,r) \leq & C\|A\|\|B\|(r_0+r)^{d-1} \\ & \times \left(\frac{1}{(1-\mu)^\alpha}\frac{t^{\alpha-d}}{r^{\alpha-d-1}} + te^{-\frac{\xi r}{t}}\right),\end{aligned} \tag{6.84}$$

where r_0 is the radius of the smallest ball that contains X and $\mu, \xi \in (0,1)$ are arbitrary constants. The second term decays exponentially with r/t and becomes negligible compared to the first term when $r \gg t$. Note that, other than its dependence on r_0, this bound is independent of the size of X, Y and is valid for $\alpha > 2d$.

Before we present the proof of Eq. (6.84), we summarize the key steps of the proof:

1. First, divide $[0,t]$ into M equal time intervals and define $t_0, t_1, \ldots, t_M$ such that $t_0 = 0$ and $t_{j+1} - t_j = \tau = t/M$. We denote by U_{t_i,t_j} the evolution unitary of the system from time t_i to t_j.

2. Setting $U_j = U_{t_{M-j},t_{M-j+1}}$ for brevity, we can decompose the evolution of A into M timesteps:

$$A(t) = U_M^\dagger U_{M-1}^\dagger \ldots U_1^\dagger A U_1 \ldots U_{M-1} U_M. \tag{6.85}$$

3. We then use a truncation technique (explicitly described below) to approximate

$U_1^\dagger A U_1$ by some operator A_1 such that

$$\|U_1^\dagger A U_1 - A_1\| = \varepsilon_1, \tag{6.86}$$

and A_1 is supported on a ball of size at most ℓ larger than the size of the support of A.

4. Repeat the above approximation for the other time slices, i.e. find $A_2, \ldots, A_M$ such that

$$\|U_2^\dagger A_1 U_2 - A_2\| = \varepsilon_2, \tag{6.87}$$

$$\|U_3^\dagger A_2 U_3 - A_3\| = \varepsilon_3, \tag{6.88}$$

$$\cdots$$

$$\|U_M^\dagger A_{M-1} U_M - A_M\| = \varepsilon_M. \tag{6.89}$$

By the end of this process, we have approximated $A(t)$ by an operator A_M whose support is at most $M\ell$ larger than the support of A.

5. By choosing $M\ell$ just smaller than r, the support of A_M does not overlap with the support of B. Therefore, $[A_M, B] = 0$, and $\mathcal{C}(t, r)$ is at most the total error of the approximation, i.e.

$$\varepsilon = \varepsilon_1 + \cdots + \varepsilon_M. \tag{6.90}$$

The total error ε, and hence the bound, depends on the truncation technique used in

Step 3. In Ref. [21], the authors used a technique inspired by digital quantum simulation, which works for $\alpha > 2d$. However, in addition to truncating the evolution unitary, the technique in Ref. [21] also truncates the Hamiltonian. The large error from this truncation makes it difficult to further improve the bound. Here, we use a different, simpler technique to generalize the bound in Ref. [21] to k-body interactions for all $k \geq 2$. Our technique does not require truncating the Hamiltonian, eliminating a hurdle for future improvements on the bound [4].

Let us start without any assumption on the interactions of the system. We only assume that there already exists a bound on the commutator norm for the system:

$$\mathcal{C}(t,r) \leq f(t,r)\phi(X)\|A\|\|B\|, \tag{6.91}$$

for some function f that increases with t and decreases with r, where $\phi(X)$ is the boundary area of X.

To truncate $U_1^\dagger A U_1$, we simply trace out the part of $U_1^\dagger A U_1$ that lies outside a ball of radius ℓ around the support of A [36]:

$$A_1 \equiv \frac{1}{\mathrm{Tr}(\mathbb{I}_{\mathcal{B}_\ell(A)^c})} \mathrm{Tr}_{\mathcal{B}_\ell(A)^c}(U_1^\dagger A U_1) \otimes \mathbb{I}_{\mathcal{B}_\ell(A)^c} \tag{6.92}$$

$$= \int_{\mathcal{B}_\ell(A)^c} d\mu(W) W (U_1^\dagger A U_1) W^\dagger, \tag{6.93}$$

where $\mathcal{B}_\ell(A)$ is a ball of radius $\ell + r_0$ centered on A and X^c denotes the complement of the

[4] We note that the approach in Ref. [21] also gives the effective Hamiltonian that generates the evolution from A to A_M, which is more useful than the technique presented here when knowing such a Hamiltonian is important, e.g. in digital quantum simulation.

set X. In Eq. (6.93), we rewrite the trace over $\mathcal{B}_\ell(A)^c$ as an integral over the unitaries W supported on $\mathcal{B}_\ell(A)^c$ and $\mu(W)$ denotes the Haar measure for the unitaries. Effectively, A_1 is the part of A that lies inside the ball $\mathcal{B}_\ell(A)$. The error from approximating $U_1^\dagger A U_1$ with A_1 is

$$\begin{aligned}
\varepsilon_1 &= \|U_1^\dagger A U_1 - A_1\| \\
&= \|U_1^\dagger A U_1 - \int_{\mathcal{B}_\ell(A)^c} d\mu(W) W (U_1^\dagger A U_1) W^\dagger\| \\
&= \| \int_{\mathcal{B}_\ell(A)^c} d\mu(W) \left[U_1^\dagger A U_1 - W (U_1^\dagger A U_1) W^\dagger \right] \| \\
&\leq \int_{\mathcal{B}_\ell(A)^c} d\mu(W) \| \left[U_1^\dagger A U_1, W \right] \|. \qquad (6.94)
\end{aligned}$$

Note that W is a unitary whose support is at least a distance ℓ from the support of A. Therefore, using the LR bound in Eq. (6.91), we have

$$\begin{aligned}
\varepsilon_1 = \|U_1^\dagger A U_1 - A_1\| &\leq \int_{\mathcal{B}_\ell(A)^c} d\mu(W) \|A\| \phi(X) f(\tau, \ell) \\
&= \|A\| \phi(X) f(\tau, \ell), \qquad (6.95)
\end{aligned}$$

where τ is the time interval of each time slice. In addition, it is clear from the definition of A_1 in Eq. (6.93) that $\|A_1\| \leq \|A\|$. Therefore, the error of the approximation in the j-th time slice is at most

$$\varepsilon_j \leq \|A\| \phi(X_{j-1}) f(\tau, \ell), \qquad (6.96)$$

where X_j is the support of A_j. Thus, the new bound is

$$\mathcal{C}(t,r) \leq 2\|B\|\varepsilon \leq 2M\|A\|\|B\|\phi_{\max} f(\tau,\ell) \tag{6.97}$$

$$= 2\|A\|\|B\|\frac{t}{\tau}\phi_{\max} f(\tau,\ell), \tag{6.98}$$

where $\phi_{\max} = \max_j \phi(X_j)$ and M has been replaced by t/τ. Note that the above bound is valid for all choices of t, ℓ, as long as

$$M = \frac{t}{\tau} < \frac{r}{\ell}, \tag{6.99}$$

$$\ell \geq 1, \tag{6.100}$$

$$\tau \leq t. \tag{6.101}$$

The first condition ensures that the operator after the last time slice A_M is still outside the support of B, while the last two are practical constraints.

Equation (6.99) is equivalent to $\ell < r\tau/t$. Because $f(\tau,\ell)$ is a decreasing function of ℓ, the bound Eq. (6.98) would be the tightest if we chose $\ell = \xi r\tau/t$ for some ξ less than, but very close to, 1. The bound Eq. (6.98) becomes

$$\mathcal{C}(t,r) \leq 2\|A\|\|B\|\phi_{\max}\, f\left(\tau, \frac{\xi r\tau}{t}\right)\frac{t}{\tau}. \tag{6.102}$$

Note that the only free parameter left is τ, which is constrained by [see Eqs. (6.99)–

(6.101)]:

$$t \geq \tau > \frac{t}{r}. \tag{6.103}$$

We are now ready to generalize the bound in Ref. [21] to many-body interactions. Plugging the k-body generalization of Eq. (6.83) [see Eq. (D.13) in Appendix D.1] into Eq. (6.102), we have

$$\begin{aligned}
\mathcal{C}(t,r) &\leq C\|A\|\|B\|\phi_{\max}\frac{t}{\tau} \\
&\quad\times\left(\frac{1}{(1-\mu)^{\alpha}}\frac{e^{v\tau}}{\left(\frac{\xi r\tau}{t}\right)^{\alpha-d-1}}+e^{v\tau-\frac{\xi r\tau}{t}}\right) \\
&\leq C\|A\|\|B\|(r_0+r)^{d-1}\frac{t}{\tau} \\
&\quad\times\left(\frac{1}{(1-\mu)^{\alpha}}\frac{e^{v\tau}}{\tau^{\alpha-d-1}}\left(\frac{t}{r}\right)^{\alpha-d-1}+e^{v\tau-\frac{\xi r\tau}{t}}\right),
\end{aligned}$$

where we have assumed without loss of generality that X is a ball of radius r_0 and replaced $\phi_{\max} \propto (r_0+r)^{d-1}$. Taking $\tau = 1$ to be a constant, we obtain a bound that is valid for all $\alpha > d+1$:

$$\begin{aligned}
\mathcal{C}(t,r) &\leq C\|A\|\|B\|(r_0+r)^{d-1} \\
&\quad\times\left(\frac{1}{(1-\mu)^{\alpha}}\frac{t^{\alpha-d}}{r^{\alpha-d-1}}+te^{-\frac{\xi r}{t}}\right).
\end{aligned}$$

In particular, if r_0 is a constant, we can simplify (in the limit of large t, r) to

$$\mathcal{C}(t,r) \leq C\|A\|\|B\|\left(\frac{1}{(1-\mu)^{\alpha}}\frac{t^{\alpha-d}}{r^{\alpha-2d}}+tr^{d-1}e^{-\frac{\xi r}{t}}\right). \tag{6.104}$$

Note that although the bound is, in principle, valid for $\alpha > d + 1$, it is only useful for $\alpha > 2d$.

6.9.2 The Heating Rate from the Linear Response Theory

In this section, we present the derivation of an exponentially suppressed heating rate for periodically driven, power-law Hamiltonians under the assumptions of linear response theory. We will assume that the drive $V(t)$ is harmonic and local. That is, we can write $V(t) = g\cos(\omega t)O$, for some small constant g and some time-independent operator $O = \sum_i O_i$ composed of local operators O_i. For simplicity, we assume each O_i is supported on a single site i (but our results also hold when O_i is supported on a finite number of sites around i). We also assume the system is initially in a thermal state ρ_β of H_0 with a temperature β^{-1}. Within the linear response theory, the energy absorption rate is proportional to the dissipative (imaginary) part of the response function $\sigma(\omega) = \sum_{i,j} \sigma_{ij}(\omega)$ [146], where

$$\sigma_{ij}(\omega) = \frac{1}{2}\int_{-\infty}^{\infty} dt e^{i\omega t} \left\langle [O_i(t), O_j(0)] \right\rangle_\beta , \tag{6.105}$$

where $\langle O \rangle_\beta \equiv \mathrm{Tr}(\rho_\beta O)$ denotes the thermal average of O, and $O(t) = e^{iH_0 t} O e^{-iH_0 t}$ is the time-evolved version of O under H_0.

The authors of Ref. [146] showed that there exists a constant κ such that for all i, j and for all $\omega, \delta\omega > 0$, the (i, j) entry of $\sigma(\omega)$ can be bounded as

$$|\sigma_{ij}([\omega, \omega + \delta\omega])| \leq e^{-\kappa\omega}, \tag{6.106}$$

where $f([\omega_1, \omega_2]) \equiv \int_{\omega_1}^{\omega_2} f(\omega) d\omega$. Although the statement of Ref. [146] applies to Hamiltonians with finite-range interactions, we show in Appendix D.2.1 that it also holds for power-law Hamiltonians for all $\alpha \geq 0$.

In principle, Eq. (6.106) already implies that the absorption rate of a finite system is exponentially small as a function of the frequency ω. However, since there are N sites in the system, naively applying Eq. (6.106) by summing over the indices i, j yields a superextensive heating rate $\sim N^2 e^{-\kappa\omega}$. Such superextensivity is non-physical, as it would imply that a local drive instigates a divergent absorption per site in the thermodynamic limit. To address this, Ref. [146] introduced a bound complementary to Eq. (6.106)—based on Lieb-Robinson bounds for finite-range interactions [12]—that implies the contribution from the off-diagonal terms is also *exponentially* suppressed with the distance r_{ij} between the sites i, j.

The case of power-law interacting Hamiltonians is somewhat more involved. Due to the long-range interaction, the commutator $\langle [O_i(t), O_j] \rangle_\beta$ can decay more slowly as a function of r_{ij} than in the finite-range case. Fortunately, we show that it still decays sufficiently quickly for us to recover the extensive, exponentially-small heating rate for power-law Hamiltonians. We provide the technical proof in Appendix D.2.2, but a high-level argument goes as follows.

Lieb-Robinson bounds for power-law systems with $\alpha > d$ [18, 21, 29, 141] imply that the contributions from the (i, j) entries are suppressed by $1/r_{ij}^\alpha$. Therefore, the total contribution to $\sigma([\omega, \omega + \delta\omega])$ from the pairs (i, j) with r_{ij} larger than some distance r_*

(to be chosen later) is at most

$$\sum_{i,j:r_{ij}\geq r_*} \frac{C}{r_{ij}^{\alpha}} \leq \frac{CN}{r_*^{\alpha-d}}, \tag{6.107}$$

where we use the same notation C to denote different constants that are independent of r_{ij}, t, and N. The factor N comes from summing over i and the factor r_*^d from summing over j at least a distance r_* from i.

For $r_{ij} \leq r_*$, we simply use the bound in Eq. (6.106) to bound their contributions:

$$\sum_{i,j:r_{ij}\leq r_*} Ce^{-\kappa\omega} \leq CNr_*^d e^{-\kappa\omega}, \tag{6.108}$$

where Nr_*^d is roughly the number of pairs (i, j) separated by distances less than r_*. Combining Eq. (6.107) with Eq. (6.108), we get $|\sigma([\omega, \omega + \delta\omega])| \leq CNr_*^d \left(e^{-\kappa\omega} + r_*^{-\alpha}\right)$. Finally, choosing $r_* = \exp(\kappa\omega/\alpha)$, we obtain a bound on the absorption rate,

$$|\sigma([\omega, \omega + \delta\omega])| \leq CN \exp\left[-\left(1 - \frac{d}{\alpha}\right)\kappa\omega\right], \tag{6.109}$$

which decays exponentially quickly with ω as long as $\alpha > d$. Thus, we have shown that, within the linear response theory, the heating rate of power-law interacting Hamiltonians obeys a bound that is qualitatively similar to that for finite-range interactions: the heating rate is extensive, but exponentially small in the driving frequency.

6.9.3 The Heating Rate from a Magnus-like Expansion

We now present a more general approach to proving a bound on the heating time in a system governed by a periodically driven, power-law Hamiltonian. In particular, this approach remains correct for strongly driven systems, where linear-response theory does not apply. We generalize Ref. [147] and construct an effective time-independent Hamiltonian H_*. The leading terms of H_* resemble the effective Hamiltonian one would get from the Magnus expansion [152, 153, 154]. Using Lieb-Robinson bounds for power-law interactions, we show that the evolution of local observables under H_* well approximates the exact evolution up to time t_*, which is exponentially long as a function of the drive frequency. Additionally, the existence of the effective Hamiltonian H_* also implies a prethermalization window during which the system could thermalize with respect to H_* before eventually heating up after time t_*.

Following Ref. [147], we construct a periodic unitary transformation $Q(t)$ such that $Q(t+T) = Q(t)$ and $Q(0) = \mathbb{I}$. After moving into the frame rotated by $Q(t)$, we show that the transformed Hamiltonian is nearly time-independent and the norm of the residual time-dependent part is exponentially small as a function of the frequency.

To construct the unitary $Q(t)$, we note that the state of the system in the rotated frame, $|\phi(t)\rangle = Q^\dagger(t)\,|\psi(t)\rangle$, obeys the Schrödinger equation with a transformed Hamiltonian $H'(t)$ ($\hbar = 1$):

$$i\partial_t\,|\phi(t)\rangle = (Q^\dagger H Q - iQ^\dagger \partial_t Q)\,|\phi(t)\rangle \equiv H'(t)\,|\phi(t)\rangle\,. \tag{6.110}$$

We write $Q = e^{\Omega}$, where $\Omega(t)$ is a periodic, anti-Hermitian operator, i.e. $\Omega(t) = \Omega(t+T)$ and $\Omega^\dagger = -\Omega$. We then assume that the period T is small so that we may expand $\Omega(t) = \sum_{q=1}^{\infty} \Omega_q$ in orders of T, where $\|\Omega_q\| = \mathcal{O}(T^q)$, and we will eventually choose Ω_q such that the transformed Hamiltonian $H'(t)$ is almost time-independent. In particular, we shall truncate the expansion of $\Omega(t)$ up to order $q_{\max}$ and choose Ω_q recursively for all $q \leq q_{\max}$ to minimize the norm of the driving term in $H'(t)$.

We can rewrite $H'(t)$ from Eq. (6.110) as:

$$H'(t) = e^{-\mathrm{ad}_\Omega}[H_0 + V(t)] - i\frac{1 - e^{-\mathrm{ad}_\Omega}}{\mathrm{ad}_\Omega}\partial_t\Omega, \tag{6.111}$$

with $\mathrm{ad}_\Omega A = [\Omega, A]$. From Eq. (6.111), we can define $H'_q(t)$ for $q = 0, 1, \ldots$ such that $H' = \sum_{q=0}^{\infty} H'_q(t)$ is expanded in powers of T:

$$H'_q(t) = G_q(t) - i\partial_t\Omega_{q+1}(t), \tag{6.112}$$

where we define G_q via $\Omega_1, \ldots, \Omega_q$ as follows:

$$\begin{aligned} G_q(t) = &\sum_{k=1}^{q} \frac{(-1)^k}{k!} \sum_{\substack{1 \leq i_1, \ldots, i_k \leq q \\ i_1 + \cdots + i_k = q}} \mathrm{ad}_{\Omega_{i_1}} \ldots \mathrm{ad}_{\Omega_{i_k}} H(t) \\ &+ i\sum_{k=1}^{q} \frac{(-1)^{k+1}}{(k+1)!} \sum_{\substack{1 \leq i_1, \ldots, i_k, m \leq q+1 \\ i_1 + \cdots + i_k + m = q+1}} \mathrm{ad}_{\Omega_{i_1}} \ldots \mathrm{ad}_{\Omega_{i_k}} \partial_t \Omega_m, \end{aligned} \tag{6.113}$$

and $G_0(t) = H(t)$. Now, recall that $\Omega_q(t)$ are operators that we can choose. From Eq. (6.112), we choose $\Omega_1(t)$ such that it cancels out the time-dependent part of $G_0(t)$, making H'_0 time-independent. This choice of $\Omega_1(t)$ also defines $G_1(t)$. We then choose $\Omega_2(t)$ to

eliminate the time-dependent part of $G_1(t)$. In general, we choose Ω_q successively from $q = 1$ to some $q = q_{\max}$ (to be specified later) so that H'_q are time-independent for all $q < q_{\max}$. Therefore, the remaining time-dependent part of the transformed Hamiltonian $H'(t)$ must be at least $\mathcal{O}(T^{q_{\max}})$. Specifically, for $q < q_{\max}$, we choose the following:

$$\bar{H}'_q = \frac{1}{T}\int_0^T G_q(t)dt, \tag{6.114}$$

$$\Omega_{q+1}(t) = -i\int_0^t \left(G_q(t') - \bar{H}'_q\right)dt'. \tag{6.115}$$

Here, Eq. (6.115) ensures that Eq. (6.112) becomes $H'_q(t) = \bar{H}'_q$, and, thus, that H'_q is time-independent for all $q < q_{\max}$. On the other hand, for $q \geq q_{\max}$, we choose $\Omega_{q+1}(t) = 0$, so that $H'_q(t) = G_q(t)$. By this construction, we can rewrite the transformed Hamiltonian into the sum of a time-independent Hamiltonian H_* and a drive $V'(t)$ that contains higher orders in T:

$$H'(t) = \sum_{q=0}^{\infty} H_q(t) = \underbrace{\sum_{q=0}^{q_{\max}-1} \bar{H'}_q}_{\equiv H_*} + \underbrace{\sum_{q=q_{\max}}^{\infty} G_q(t)}_{\equiv V'(t)}, \tag{6.116}$$

As a result of the transformation, the driving term $V'(t)$ is now $\mathcal{O}(T^{q_{\max}})$. As discussed before, we will eventually choose the cutoff $q_{\max}$ to minimize the norm of the residual drive $\|V'(t)\|$.

To estimate the norm of $V'(t)$, elucidating its dependence on $q_{\max}$, we first need more information on the structure of the $\Omega_q(t)$ for all $1 \leq q \leq q_{\max}$. In particular, we show that G_q and Ω_q are power-law interacting Hamiltonians. To do so, we first need to define

some more notation. We denote by $\mathbb{H}_\alpha$ the set of power-law Hamiltonians with exponent α and a local energy scale $\eta = 1$. In addition, we denote by $\mathbb{H}_\alpha^{(k)}$ the subset of $\mathbb{H}_\alpha$ which contains all power-law Hamiltonians whose local support size [see Definition 1] is at most $k+1$. For a real positive constant a, we also denote by $a\mathbb{H}_\alpha$ the set of Hamiltonians H such that $a^{-1}H$ is a power-law Hamiltonian with the same exponent α.

The following lemma says that G_q and Ω_q are also power-law Hamiltonians up to a prefactor.

Lemma 7. *For all $q < q_{max}$, we have*

$$G_q \in T^q q! c^q \lambda^q \mathbb{H}_\alpha^{(q+1)}, \tag{6.117}$$

$$\partial_t \Omega_{q+1} \in T^q q! c^q \lambda^q \mathbb{H}_\alpha^{(q+1)}, \tag{6.118}$$

$$\Omega_{q+1} \in T^{q+1} q! c^q \lambda^q \mathbb{H}_\alpha^{(q+1)}, \tag{6.119}$$

where c, λ are constants to be defined later.

Observe that for any order q, the last two bounds, i.e. Eq. (6.118) and Eq. (6.119), follow immediately from Eq. (6.117) and the definition of Ω_q. Note that Lemma 7 holds for $G_0(t) = H(t) \in \mathbb{H}_\alpha^{(1)}$. It is also straightforward to prove Lemma 7 inductively on q. The factor T^q comes from the constraint in Eq. (6.113) that $i_1 + \cdots + i_k = q$, along with the fact that each Ω_{i_ν} is $\mathcal{O}(T^{i_\nu})$ for all $\nu = 1, \ldots, k$. Similarly, the factor of $q!$ is combinatorial and comes from the nested commutators in Eq. (6.113). We provide a more technical proof of Lemma 7 in Appendix D.3.

As a consequence of Lemma 7, we can bound the local norms of the operators:

$$\|G_q\|_l \leq T^q q! c^q \lambda^{q+1} \leq \lambda e \sqrt{q} \left(\frac{Tqc\lambda}{e} \right)^q, \tag{6.120}$$

$$\|\Omega_q\|_l \leq T^q (q-1)! c^{q-1} \lambda^q \leq \frac{e}{c} \left(\frac{Tqc\lambda}{e} \right)^q. \tag{6.121}$$

There are two competing factors in the bounds: T^q, which decreases with q, and $q! \sim q^q$, which increases with q. This suggests that the optimal choice for $q_{\max}$—in order to minimize the local norm in Eq. (6.120)—should be around $e/(cT\lambda)$. In the following, we shall choose

$$q_{\max} = \omega_* \equiv \frac{e}{cT\lambda} e^{-\kappa}, \tag{6.122}$$

for some $\kappa > \ln 2$. Note also that ω_* is equal to frequency $\omega = 1/T$ up to a constant. With this choice of $q_{\max}$, Eq. (6.117) reduces to

$$G_q \in \lambda e \sqrt{q} e^{-\kappa q} \mathbb{H}_\alpha^{(q+1)}, \tag{6.123}$$

for all $q < q_{\max} = \omega_*$. By summing over G_q with $q < \omega_*$, we find that the effective time-independent Hamiltonian H_* [see Eq. (6.116)] is also a power-law Hamiltonian, i.e. $H_* \in C\mathbb{H}_\alpha^{(q_{\max})} \in C\mathbb{H}_\alpha$, up to a constant C that may depend only on κ.

Similarly, we find from Eq. (6.119) that $\Omega_q \in e/(c\lambda) e^{-\kappa q} \mathbb{H}_\alpha^{(q)}$ for all $q \leq \omega_*$. Plugging into the definition of G_q and noting that we choose $\Omega_q = 0$ for all $q \geq q_{\max}$, we

find an identity similar to Eq. (6.123), but for $q \geq \omega_*$:

$$G_q \in Ce^{-\kappa' q}\mathbb{H}_\alpha, \tag{6.124}$$

where $\kappa' > \kappa - \ln 2$ is a constant. Summing over G_q with $q \geq q_{\max}$ [see Eq. (6.116)], we again find that the residual drive $V'(t)$ is a power-law Hamiltonian up to a prefactor that decays exponentially with ω_*:

$$V'(t) \in Ce^{-\kappa'\omega_*}\mathbb{H}_\alpha, \tag{6.125}$$

where C and κ' are some positive constants. As a result, the local norm of $V'(t)$ decreases exponentially with ω_*: $\|V'(t)\|_l \leq C\lambda e^{-\kappa'\omega_*}$.

As discussed earlier, Eq. (6.116) and Eq. (6.125) imply the existence of an effective time-independent Hamiltonian H_* such that the difference $\|Q^\dagger HQ - H_*\| = \|V'\|$ is exponentially small as a function of $\omega_* \propto 1/T$. However, even if $\|V'\|_l$ is exponentially small, $\|V'\|$ still diverges in the thermodynamic limit. Therefore, in order to characterize the heating rate of the Hamiltonian, it is necessary to investigate the evolution of a local observable O under $H(t)$. We show that the evolution is well described by the effective time-independent Hamiltonian H_* at stroboscopic times $t = T\mathbb{Z}$. Without loss of generality, we assume the local observable O is supported on a single site and $\|O\| = 1$. Following a similar technique used in Abanin et al. [147], we write the difference between the approximate evolution under the effective Hamiltonian and the exact evolution (in the

rotated frame):

$$\begin{aligned}\delta &= Q(t)U^\dagger(t)OU(t)Q^\dagger(t) - e^{itH_*}Oe^{-itH_*} \\ &= i\int_0^t ds W^\dagger(s,t)\left[V'(s), e^{isH_*}Oe^{-isH_*}\right]W(s,t),\end{aligned}$$

where $U(t) = \mathcal{T}\exp\left(-i\int_0^t H(t')dt'\right)$ is the time evolution generated by the full Hamiltonian $H(t)$ and $W(s,t) = \mathcal{T}\exp\left(-i\int_s^t H'(t')dt'\right)$ is the evolution from time s to t generated by $H'(t)$. We can then bound the norm of the difference using the triangle inequality:

$$\|\delta\| \le \int_0^t ds \|\left[V'(s), e^{isH_*}Oe^{-isH_*}\right]\|. \tag{6.126}$$

We can bound the right-hand side of Eq. (6.126) using Lieb-Robinson bounds for power-law interactions.

First, we provide an intuitive explanation why the norm of δ is small for small enough time ($t \ll N$). Recall that the operator O is initially localized on a single site. At small time, it is still quasilocal and therefore significantly noncommutative with only a small number of terms of V' lying inside the "light cone" generated by the evolution under H_*. There are several Lieb-Robinson bounds for power-law interactions [19, 21, 29, 141] [see also Eq. (D.13) and Eq. (6.84)], each provides a different estimate for the shape of the light cone, resulting in a different bound for the heating time.

If the light cone is logarithmic (as bounded in Ref. [29]), the commutator norm in Eq. (6.126) would grow exponentially quickly with time and eventually negate the

exponentially-small factor $\exp(-\kappa'\omega_*)$ from $\|V'\|_l$. Therefore, in such cases, the system could potentially heat up only after $t_* \propto \omega_* = 1/T$. On the other hand, if we use the Lieb-Robinson bounds that imply algebraic light cones (as in Refs. [19, 21, 141] for $\alpha > 2d$), the commutator norm only grows subexponentially with time, and we can expect to recover the exponentially-long heating time $t_* \propto e^{\kappa'\omega_*}$ derived for finite-range interactions [147, 148].

Appendix D.4 contains the mathematical details, but the results of this analysis are as follows. Using Gong et al. [29] [or its k-body generalization Eq. (D.12)], which holds for $\alpha > d$ and has a logarithmic light cone $t \gtrsim \log r$, yields:

$$\|\delta\| \leq Ce^{-\kappa'\omega_*}e^{2dvt/\alpha}. \tag{6.127}$$

Thus, the difference δ is only small for time $t_* \propto \omega_* \propto 1/T$. This behavior is expected because the region inside the light cone implied by Gong et al. 's bound expands exponentially quickly with time.

If instead we use the bound in Else et al. [141], we find:

$$\|\delta\| \leq Ce^{-\kappa'\omega_*}\xi\left(\frac{d}{1-\sigma}\right)t^{\frac{d}{1-\sigma}+1}, \tag{6.128}$$

where $\xi(x) \equiv \frac{1}{x}2^x\Gamma(x)$ and Γ is the Gamma function. Thus, the difference is small up to an exponentially long time $t_* \propto e^{\kappa'\omega_*\frac{1-\sigma}{d+1-\sigma}}$. The result holds for $\alpha > d\left(1+\frac{1}{\sigma}\right)$, where σ can be chosen arbitrarily close to 1. This condition is effectively equivalent to $\alpha > 2d$ [see Appendix D.4 for a discussion of the limit $\sigma \to 1^-$].

We may also use the bound in Tran et al. [21] [see Eq. (6.104) for its generalization to k-body interactions], which gives

$$\|\delta\| \leq Ce^{-\kappa'\omega_* t^{\frac{d(\alpha-d)}{\alpha-2d}+1}}. \tag{6.129}$$

Thus, the difference is small up to an exponentially-long time $t_* \propto \exp\left(\kappa'\omega_* \frac{\alpha-2d}{\alpha(d+1)-d(d+2)}\right)$. This analysis works only when $\alpha > 3d$, but, within this regime, the exponent of the heating time using this bound is larger than obtained in Eq. (6.128). This is due to the trade-off between the tail and the light cone between the bounds in Refs. [21, 141]. See Appendix D.4 for more details.

Finally, we conjecture a tight bound for power-law interactions that holds for all $\alpha > d$, and we will provide the full derivation of δ for such a bound. First, we consider the light cone of such a bound. Given the best known protocols for quantum information transfer [24], the best light cone we could hope for would be $t \gtrsim r^{\alpha-d}$ for $d+1 > \alpha > d$ and linear for $\alpha > d+1$. In the following, we assume the light cone of the conjectured bound is $t \gtrsim r^{1/\beta}$ for some constant $\beta \geq 1$ for all $\alpha > d$.

Next, we consider the tail of the bound, i.e. how the conjectured bound decays with the distance at a fixed time. Since it is always possible to signal between two sites using their direct interaction, which is of strength $1/r^\alpha$, the tail of the bound cannot decay faster than $1/r^\alpha$. We shall assume that the bound decays with the distance exactly as $1/r^\alpha$.

For simplicity, we assume that the conjectured bound takes the form

$$\|[A(t), B]\| \leq C\|A\|\|B\| \left(\frac{t^\beta}{r}\right)^\alpha, \tag{6.130}$$

which manifestly has a light cone $t \gtrsim r^{1/\beta}$ and decays as $1/r^\alpha$ with the distance. Let $r_*(t) = t^\beta$ be the light cone boundary and consider the sum inside and outside the light cone.

For convenience, denote $V'' = C^{-1}e^{\kappa'\omega_*}V, \bar{H}'' = \gamma^{-1}\bar{H}'$ so that $V'', \bar{H}'' \in \mathbb{H}_\alpha$. Since $V''(s)$ is a power-law Hamiltonian, it follows that $\|V''_r(t)\| \leq Cr^{d-1}$. We can rewrite the bound on $\|\delta\|$ as

$$\|\delta\| \leq Ce^{-\kappa'\omega_*} \int_0^t ds \| \Big[V''(s), e^{is\gamma\bar{H}''} O e^{-is\gamma\bar{H}''} \Big] \|, \tag{6.131}$$

Now write $V''(s) = \sum_{r=0}^\infty V''_r(s)$, where $V''_r(s) \equiv \sum_{X:\text{dist}(X,O)\in[r,r+1)} h_X$ denotes the terms of $V''(s)$ supported on subsets exactly a distance between r and $r+1$ away from O. Writing the sum this way, we can now separate terms inside and outside of the light cone.

For the terms inside the light cone, we bound:

$$\begin{aligned} &\sum_{r\leq r_*(s)} \| \Big[V''_r(s), e^{is\gamma\bar{H}''} O e^{-is\gamma\bar{H}''} \Big] \| \\ &\leq 2 \sum_{r\leq r_*(s)} \|V''_r(s)\|\|O\| \leq Cr_*(s)^d \leq Cs^{\beta d}. \end{aligned} \tag{6.132}$$

For the terms outside the light cone, we use the conjectured bound:

$$\sum_{r>r_*(s)} \|\left[V_r''(s), e^{is\gamma\bar{H}''}Oe^{-is\gamma\bar{H}''}\right]\|$$

$$\leq C \sum_{r>r_*(s)} \|V_r''(s)\|\|O\|\frac{s^{\beta\alpha}}{r^\alpha} \tag{6.133}$$

$$\leq C \sum_{r>r_*(s)} \frac{s^{\beta\alpha}}{r^{\alpha-d+1}} \leq C\frac{s^{\beta\alpha}}{r_*(s)^{\alpha-d}} = Cs^{\beta d}. \tag{6.134}$$

Combining Eq. (6.132) and Eq. (6.134), we get

$$\|\delta\| \leq Ce^{-\kappa'\omega_*}t^{\beta d+1}, \tag{6.135}$$

which implies an exponential heating time as a function of ω_*, i.e. $t_* \propto \exp(\kappa'\omega_*/(\beta d+1))$. Recall that the best values we can hope for β are $\beta = 1/(\alpha-d)$ when $d+1 > \alpha > d$ and $\beta = 1$ when $\alpha > d+1$. Note also that the exponential heating time would hold for all $\alpha > d$, matching the result given by the linear response theory.

6.9.4 Discussion

Our work generalizes the results of Refs. [146, 147, 148] for finite-range interactions to power-law interactions. Using two independent approaches, we show that periodically driven, power-law systems with a large enough exponent α can only heat up after time that is exponentially long in the drive frequency. The results only hold if α is larger than some critical value α_c. Physically, the existence of α_c coincides with our expectation that power-law interactions with a large enough exponent α are effectively short-range.

However, the two approaches imply different values for α_c. While both the Magnus expansion in Ref. [137] and the Magnus-like expansion in this section independently suggest $\alpha_c = 2d$, the linear response theory implies $\alpha_c = d$. We conjecture that this gap is due to the lack of tighter Lieb-Robinson bounds for power-law interactions, especially for α between d and $2d$. Indeed, we demonstrated in Sec. 6.9.3 that a tight Lieb-Robinson bound for this range of α implies an exponentially-long heating time for all $\alpha > d$, matching the result from the linear response approach, as well as previous numerical evidence for systems with $\alpha < 2d$ [150]. Therefore, proving a tight Lieb-Robinson bound has important implications for the heating time of power-law interacting systems.

Part II

Symmetry and Digital Simulation of Quantum Systems

Chapter 7: Destructive Error Interference in Product-Formula Lattice Simulation

In this chapter, we provide an approach to tighten the error bound of the first-order product formula (PF1) for simulating several physically relevant systems, including those with nearest-neighbor interactions. Despite its simplicity and wide applicability, a tight error bound for PF1 in simulating many physically relevant systems remains elusive. Recent works [2, 3, 4] estimated that $\mathcal{O}\left(n^2t^2\right)$ elementary gates suffice to simulate the dynamics of a nearest-neighbor interacting system consisting of n sites for time t using PF1 [1]. However, the numerical evidence in Ref. [7] suggests that PF1 performs much better than this in practice. In particular, the gate count for simulating the dynamics of a nearest-neighbor Heisenberg spin chain of length n for time $t = n$ scales only as $\mathcal{O}\left(n^{2.964}\right)$. In addition, Heyl et al. [155] also found that the error of simulating the time evolution of a local observable using PF1 can be much smaller than theoretically estimated.

The key message of the chapter is that the errors from different steps of the algorithm can combine destructively, resulting in a smaller total error than previous analysis

[1] Refs. [2, 3, 4] took into account the commutativity between some interaction terms in the Hamiltonian of a nearest-neighbor interacting system. Without this commutativity, the gate count would be $\mathcal{O}\left(n^3t^2\right)$ [5, 6]

estimates. In particular, the tighter error bound suggests that simulating the dynamics of a nearest-neighbor interacting system of n sites for time t up to an error tolerance ε requires only $\max\left\{\mathcal{O}\left(n^2 t/\varepsilon\right), \mathcal{O}\left(n^2 t^{3/2}/\varepsilon^{1/2}\right)\right\}$ quantum gates, which is asymptotically smaller than the state-of-the-art bound $\mathcal{O}\left(n^2 t^2/\varepsilon\right)$ in Refs. [2, 3, 4]. At $t = n$ and at a fixed ε, our estimate $\mathcal{O}\left(n^3\right)$ also closely matches the empirical gate count $\mathcal{O}\left(n^{2.964}\right)$ computed in Ref. [7].

7.1 Setup

We assume that the system evolves under a Hamiltonian $H = \sum_X h_X$, which is a sum of time-independent terms h_X, each acting nontrivially on a subset X of constant size. Our approach applies if there exists a partition $H = H_1 + H_2$ such that the terms h_X in H_1 mutually commute and the terms h_X in H_2 also mutually commute. Examples of Hamiltonians that satisfy this assumption include all one-dimensional, finite-range [2] interacting systems, such as the Heisenberg model and the transverse field Ising model in one dimension with either open or periodic boundary conditions, and with or without disorder. Additionally, this assumption also covers some physically relevant systems in higher dimensions, such as the transverse field Ising model with either finite-range or long-range interactions.

To simulate the time-evolution of the system for time t using elementary quantum

[2] For interactions of maximum range R, we can group $\lceil (R+1)/2 \rceil$ consecutive sites into distinct blocks such that the Hamiltonian consists of only interactions between nearest-neighbor blocks. The error analysis for using the first-order product formula to simulate such a system would follow from our analysis for simulating nearest-neighbor interactions. Note, however, that we assume that the exact simulation of the evolution of each constant-size block requires only a constant amount of elementary gates.

gates, we use the first-order product formula [1]:

$$U_t \approx \left[U_{t/r}^{(1)} U_{t/r}^{(2)}\right]^r, \tag{7.1}$$

where $U_t \coloneqq \exp(-iHt)$, $U_{t/r}^{(p)} \coloneqq \exp(-iH_p t/r)$ for $p = 1, 2$, and r is the number of time segments to be chosen later so that the norm of the total error $\Delta \coloneqq U_t - [U_{t/r}^{(1)} U_{t/r}^{(2)}]^r$ is at most a constant ε. By our assumption that the terms within H_p ($p = 1, 2$) mutually commute, we can further decompose the evolution $U_{t/r}^{(p)}$ into a product of elementary quantum gates with no additional error.

For simplicity, we demonstrate our approach to estimating the gate count of PF1 on a one-dimensional lattice of n sites, evolving under a time-independent, nearest-neighbor Hamiltonian $H = \sum_{i=1}^{n-1} h_i$, where h_i is supported only on sites $i, i+1$, $\|h_i\| \leq J$ for all i, J is a constant, and $\|\cdot\|$ denotes the operator norm. Without loss of generality, we also assume $J = 1$, which sets the time scale for the dynamics of the system. We then apply PF1 to the partition $H = H_1 + H_2$, where $H_1 = \sum_{\text{odd } j} h_j$ and $H_2 = \sum_{\text{even } j} h_j$. Note that the terms within H_1 (H_2) mutually commute and therefore satisfy the aforementioned assumption.

7.2 Leading Error Contributions

To estimate the gate count, we first need a bound on the total error Δ. The previous best bound from Ref. [4] gives $\|\Delta\| \leq \mathcal{O}\left(nt^2/r\right)$, so that $r = \Theta\left(nt^2/\varepsilon\right)$ suffices to ensure the total error at most ε, giving gate count $nr = \mathcal{O}\left(t^2 n^2/\varepsilon\right)$. Before we prove our tighter bound, we will first argue simply based on the lowest order error that $\|\Delta\| \approx \mathcal{O}\left(nt/r\right)$,

which would result in a gate count $\mathcal{O}\left(tn^2/\varepsilon\right)$, matching the empirical estimate of about $\mathcal{O}\left(n^3\right)$ for $t = n$ in Ref. [7].

Let $\delta = U_{t/r} - U_{t/r}^{(1)} U_{t/r}^{(2)}$ be the error of the approximation in each time segment. In the limit $r \gg t$, the leading contribution to δ is given by the commutator between H_1 and H_2 [1]:

$$\|\delta\| \approx \frac{1}{2}\frac{t^2}{r^2}\|[H_1, H_2]\| = \mathcal{O}\left(\frac{nt^2}{r^2}\right). \tag{7.2}$$

Replacing $U_{t/r}^{(1)} U_{t/r}^{(2)}$ by $U_{t/r} + \delta$ on the right-hand side of Eq. (7.1) and expanding to first order in δ, we have an approximation for the total error:

$$\Delta \approx \sum_{j=0}^{r-1} U_{t/r}^{j}\ \delta\ U_{t/r}^{r-1-j} = \left(\sum_{j=0}^{r-1} U_{t/r}^{j}\ \delta\ U_{t/r}^{-j}\right) U_{t/r}^{r-1}, \tag{7.3}$$

where $U_{t/r}^{j} \coloneqq (U_{t/r})^j$. If we bound $\|\Delta\|$ using the triangle inequality, i.e.,

$$\|\Delta\| \approx \left\|\sum_{j=0}^{r-1} U_{t/r}^{j}\ \delta\ U_{t/r}^{r-1-j}\right\| \leq r\|\delta\| \approx \mathcal{O}\left(\frac{nt^2}{r}\right), \tag{7.4}$$

we get the same error bound (and hence the same gate count) as Ref. [4].

To understand the key idea for improving the bound, imagine δ as a vector in the space of operators and the unitary evolution $U_{t/r}^{j}\delta U_{t/r}^{-j}$ as a rotation of the vector by a small angle proportional to jt/r. Since the terms in Eq. (7.3) correspond to rotations of δ by evenly spaced angles [3], the sum involves significant cancellation, making it much

[3] While it is common in practice to divide the evolution into even time slices, our interpretation of Eq. (7.3) as a sum of the same vector δ rotated by different angles suggests that we can also achieve destructive error interference even if the slices are uneven.

smaller than the upper bound derived using the triangle inequality [Eq. (7.4)].

To realize this intuition, we make a change of variables to $x = tj/r$ and approximate the sum in Δ by an integral:

$$\|\Delta\| \approx \left\| \sum_{j=0}^{r-1} U_{t/r}^{j} \, \delta \, U_{t/r}^{-j} \right\| \approx \frac{r}{t} \left\| \int_0^t dx U_x \, \delta \, U_{-x} \right\|. \tag{7.5}$$

With the assumption that $H = H_1 + H_2$ is a sum of two terms, we rewrite δ (to leading order in t/r) as

$$\delta \approx \frac{1}{2} \left[H_1, H_2 \right] \frac{t^2}{r^2} = \frac{1}{2} \left[H, H_2 \right] \frac{t^2}{r^2}, \tag{7.6}$$

and use the identity

$$U_t \, A \, U_{-t} - A = -i \int_0^t dx U_x \left[H, A \right] U_{-x}, \tag{7.7}$$

with $A = \frac{t^2}{2r^2} H_2$, to evaluate the integral in Eq. (7.5) and arrive at an estimate for the norm of Δ:

$$\begin{aligned} \|\Delta\| &\approx \frac{r}{t} \left\| \int_0^t dx U_x \left[H, \frac{t^2}{2r^2} H_2 \right] U_{-x} \right\| \\ &\leq \frac{t}{2r} 2 \|H_2\| = \mathcal{O}\left(\frac{nt}{r} \right), \end{aligned} \tag{7.8}$$

which is a factor of t tighter than Eq. (7.4). We attribute this improvement to the *destructive interference* between the error terms in Eq. (7.3). To ensure that the total error $\|\Delta\|$ is at most ε, we choose $r = \Theta\left(nt/\varepsilon\right)$, leading to the total gate count $\mathcal{O}\left(nr\right) = \mathcal{O}\left(n^2 t/\varepsilon\right)$,

which has optimal scaling in t [156]. At $t = n$ and fixed ε, the gate count becomes $\mathcal{O}(n^3)$, which closely matches the empirical performance $\mathcal{O}(n^{2.964})$ observed in Ref. [7].

Additionally, if the time step $t/r = \tau$ is a constant, the total error of the simulation $\|\Delta\| = \mathcal{O}(n\tau)$ appears to be independent of the total number of time segments. This feature agrees well with Ref. [155], where the authors argue that for a fixed, small value of τ, the error in simulating the evolution of a local observable using PF1 would not increase with the total simulation time t. However, our bound is more general; it applies to the error in simulating the evolution unitary of the system, and hence any observable.

7.3 Higher-Order Error Contributions

We made three approximations in deriving Eq. (7.8). First, in Eq. (7.6), we considered δ to only the leading order in t/r and discarded terms of higher order in t/r. We then expanded Δ in Eq. (7.3) to only the first order in δ while ignoring the higher-order terms in δ^k. Additionally, we evaluated the sum in Eq. (7.5) by approximating it with an integral. We now make the estimation rigorous by considering the errors incurred upon making the three approximations.

First, we show that higher-order terms in t/r in the expansion of δ are indeed dominated by the second order. For that, we write δ as a series in t/r:

$$\delta := U_{t/r} - U_{t/r}^{(1)} U_{t/r}^{(2)} = \sum_{k=2}^{\infty} \frac{(-it)^k}{k! r^k} \delta_k, \tag{7.9}$$

where δ_k are operators independent of t, r. If we only need a bound on the norm of δ, it is sufficient to bound the norms of δ_k. However, in addition to the norm, we are also

interested in the structure of δ_k, described in Lemma 8, which is crucial for evaluating the total error [See Eq. (7.6)].

Lemma 8. *For all $k \geq 2$, there exist S_k, V_k such that $\delta_k = [H, S_k] + V_k$ and*

$$\|V_k\| = \mathcal{O}\left(e^{k-2} n^{k-2}\right), \tag{7.10}$$

$$\|S_k\| = \mathcal{O}\left(k^2 n^{k-1}\right), \tag{7.11}$$

$$\|[H, S_k]\| = \mathcal{O}\left(k^3 n^{k-1}\right), \tag{7.12}$$

where the big-O constants do not depend on k.

Lemma 8 holds for $k = 2$, with $S_2 = H_2$ and $V_2 = 0$ [See Eq. (7.6)]. For $k > 2$, we construct S_k, V_k inductively using the definition of δ_k in Eq. (7.9). The factor n^{k-2} in the norm of V_k comes from the $(k-2)$-th nested commutators in the expansion of δ_k. We provide a detailed proof of the lemma in Appendix E.

A corollary of Lemma 8 is $\|\delta_k\| = \mathcal{O}\left(e^k n^{k-1}\right)$, and therefore, we can immediately bound the norm of δ:

$$\begin{aligned} \|\delta\| &\leq \sum_{k=2}^{\infty} \frac{t^k}{k! r^k} \|\delta_k\| = \mathcal{O}\left(\frac{nt^2}{r^2} \sum_{k=0}^{\infty} \frac{(ent)^k}{k! r^k}\right) \\ &= \mathcal{O}\left(\frac{nt^2}{r^2} \exp \frac{ent}{r}\right) = \mathcal{O}\left(\frac{nt^2}{r^2}\right), \end{aligned} \tag{7.13}$$

where we assume $r > ent$. We later fulfill this condition by choosing an appropriate value for r.

Another corollary of Lemma 8 is that $\delta = [H, S] + V$, where $S = \sum_{k=2}^{\infty} \frac{(-it)^k}{k! r^k} S_k$

and $V = \sum_{k=3}^{\infty} \frac{(-it)^k}{k! r^k} V_k$. It is straightforward to verify the bounds on the norms of S and V:

$$\|S\| = \mathcal{O}\left(\frac{nt^2}{r^2}\right), \quad \|V\| = \mathcal{O}\left(\frac{nt^3}{r^3}\right), \tag{7.14}$$

where we again assume $r > ent$.

Next, we rectify the approximation in Eq. (7.5) by rigorously bounding the norm of the sum.

Lemma 9. *For any positive integer* $a \geq 1$,

$$\left\| \sum_{j=0}^{a-1} U_{t/r}^{j} \delta \, U_{t/r}^{-j} \right\| = \mathcal{O}\left(\frac{nt}{r}\right) + \mathcal{O}\left(a\frac{nt^3}{r^3}\right). \tag{7.15}$$

When $a = r$, the left-hand side of Eq. (7.15) is exactly the sum in Eq. (7.5). We bound the sum by approximating it with an integral, which yields $\mathcal{O}\left(nt/r\right)$ after evaluation. Carefully bounding the error of the approximation results in the second term $\mathcal{O}\left(ant^3/r^3\right)$. We present the detailed proof of the lemma in Appendix E.

Given Lemma 8 and Lemma 9, we now bound the total error $\|\Delta\|$. We expand Δ as a series in δ and write $\Delta = \sum_{k=1}^{r} \Delta_k$, where Δ_k involves only the k-th order in δ. For example, $\Delta_1 = \sum_{j=0}^{r-1} U_{t/r}^{j} \delta \, U_{t/r}^{-j}$, the norm of which we can already bound using Lemma 9. We can use the same technique to estimate $\|\Delta_k\|$ for all $k \geq 1$ (Appendix E):

$$\|\Delta_k\| \leq r^{k-1} \|\delta\|^{k-1} \mathcal{O}\left(\frac{nt}{r} + \frac{nt^3}{r^2}\right). \tag{7.16}$$

Finally, we bound $\|\Delta\|$ using the triangle inequality:

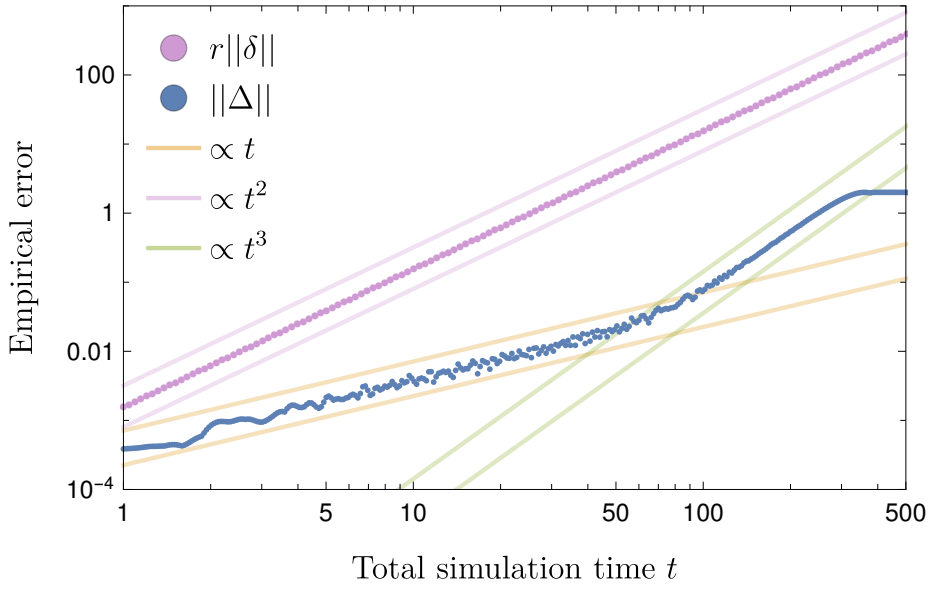

Figure 7.1: The total error $\|\Delta\|$ (blue dots) of PF1 in simulating the Heisenberg chain in Eq. (7.18) is numerically evaluated at $n = 8$, $r = 10000$, and variable time t between 0 and 1000. The purple dots represent the error estimate $r\|\delta\|$ one would get using the triangle inequalities [Eq. (7.4)]. We also plot functions proportional to t (orange lines), t^2 (purple lines), and t^3 (green lines) for reference.

$$\|\Delta\| \leq \sum_{k=1}^{r} \|\Delta_k\| = \mathcal{O}\left(\frac{nt}{r} + \frac{nt^3}{r^2}\right), \tag{7.17}$$

where we assume $r\,\|\delta\| < 1/2$ so that $\sum_{k=1}^{r} (r\|\delta\|)^{k-1} = \mathcal{O}(1)$. With our choice of r, this assumption later reduces to $\varepsilon t \leq 1$, where ε is the error tolerance of the simulation.

7.4 Empirical Error Scaling

We now benchmark the bound in Eq. (7.17) against the empirical error in simulating the dynamics of a nearest-neighbor Heisenberg chain:

$$H = \sum_{i=1}^{n-1} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1}, \tag{7.18}$$

where $\vec{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ denotes the Pauli matrices on qubit i. Using fixed values for n and r, we compute the total error of PF1 at different times t and plot the result in Fig. 7.1. We also plot in Fig. 7.2 the empirical errors of simulating the same system using the second-order (PF2) and the fourth-order (PF4) product formulae [6].

From Fig. 7.1, the total error of PF1 appears to agree well with our bound in Eq. (7.17). The change in the error scaling from $O(t)$ at small time to $O(t^3)$ at large time can be explained by the destructive error interference between the time slices as follows. While the leading error terms in each time slice scale as $O(t^2)$, they interfere destructively between time slices, resulting in a total contribution that increases with time at a slower rate $O(t)$ [recall Eq. (7.8)]. Meanwhile, some higher-order error terms do not interfere destructively. They scale as $O(t^3)$ and eventually take over as the primary contribution to the total error. This intuition also explains the similarity between the error scalings of PF1 (at late time) and PF2 [Fig. 7.2]. On the other hand, if there were no destructive error interference between the time slices, the contribution from the leading error terms to the total error of PF1 would have scaled as $O(t^2)$ [Fig. 7.1, purple dots] and saturated at 2 before the higher-order terms could take over.

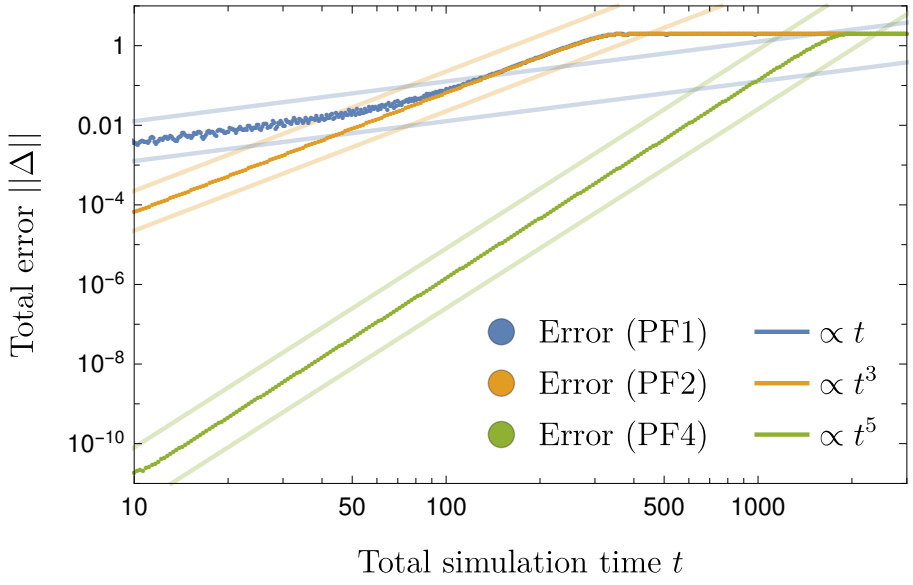

Figure 7.2: The total error of simulating the Heisenberg chain with $n = 8$ spins in Eq. (7.18) using PF1 (blue dots), PF2 (orange dots) and PF4 (green dots) is numerically computed at $r = 10000$, and variable time t between 10 and 3000. We also plot functions proportional to t (blue lines), t^3 (orange lines), and t^5 (green lines) for reference.

We also note that the error of PF2 [PF4] scales as t^3 [t^5] initially before saturating at a later time, in agreement with the existing bounds using the triangle inequality for the higher-order product formulae [4, 7]. This suggests the absence of significant destructive error interference for PF2 and PF4 in our numerical simulation.

7.5 Estimation of the Gate Count

Given the error bound in Eq. (7.17), we now count the number of gates for PF1. Equation (7.17) suggests we should choose

$$r \propto \max\left\{\frac{nt}{\varepsilon}, \sqrt{\frac{nt^3}{\varepsilon}}, 1\right\}, \tag{7.19}$$

so that the total error $\|\Delta\|$ is at most ε. First, we assume $nt \geq \varepsilon$ and consider two cases, corresponding to $\varepsilon t \leq 1$ (small time) and $\varepsilon t > 1$ (large time). The former condition implies that the first term in Eq. (7.19) dominates and therefore we should choose $r = \Theta\left(nt/\varepsilon\right)$. This choice of r together with $\varepsilon t \leq 1$ also fulfills the condition $r\|\delta\| < 1/2$ required earlier, as long as we choose a large enough prefactor in $\Theta\left(nt/\varepsilon\right)$. Thus, when $\varepsilon t \leq 1$, the gate count of PF1 is

$$\mathcal{O}\left(rn\right) = \mathcal{O}\left(\frac{n^2 t}{\varepsilon}\right). \tag{7.20}$$

On the other hand, when $\varepsilon t > 1$, we divide the simulation into m stages. In each stage, we simulate the evolution for time t/m with an error at most ε/m by further dividing the stage into r time segments. In order to apply the above analysis in each stage, we require m to be large enough so that $\varepsilon t/m^2 \leq 1$. Since the resulting gate count $\mathcal{O}\left(mn^2 t/\varepsilon\right)$ increases with m, it is optimal to choose m as small as possible, i.e. $m = \lceil\sqrt{\varepsilon t}\rceil$. Therefore, the total gate count in this case is

$$\mathcal{O}\left(\sqrt{\varepsilon t}\frac{n^2 t}{\varepsilon}\right) = \mathcal{O}\left(\frac{n^2 t^{3/2}}{\varepsilon^{1/2}}\right). \tag{7.21}$$

Finally, when $nt < \varepsilon$, we simply choose $r = \Theta(1)$, giving gate count $\mathcal{O}\left(1\right)$. Combining the above arguments, we have an upper bound on the total gate count of

$$\max\left\{\mathcal{O}\left(\frac{n^2 t}{\varepsilon}\right), \mathcal{O}\left(\frac{n^2 t^{3/2}}{\varepsilon^{1/2}}\right), \mathcal{O}\left(1\right)\right\}, \tag{7.22}$$

which is valid for all times t and is tighter than the previous best estimate in Ref. [4].

7.6 Discussion

As mentioned earlier, we assume that the terms of the Hamiltonian can be separated into two parts such that the terms within each part mutually commute. Therefore, our results apply to translationally invariant spin chains in one dimension with finite-range interactions and with either open or periodic boundary conditions, as well as disordered spin chains, such as those featuring many-body localization [157]. Additionally, our analysis also holds for some systems in higher dimensions, such as the transverse field Ising model with either finite-range or long-range interactions, where the two mutually commuting parts of the Hamiltonian are the spin-spin interactions and the field terms. However, for long-range interactions, the number of interaction terms can scale as $\mathcal{O}\left(n^2\right)$ [instead of $\mathcal{O}\left(n\right)$ for the finite-range interactions], so the scalings of the error bound and of the gate count as functions of n must be adjusted accordingly. Furthermore, our technique can also be used to bound the error in simulating materials where the electronic structure Hamiltonian in the plane wave dual basis [158] is a sum of mutually commuting kinetic energy terms and Coulomb interactions.

However, it is unclear whether our approach generalizes to Hamiltonians that can only be separated into three or more mutually commuting parts, such as those that typically occur in higher dimensions and systems with general long-range interactions, where the simple relation between δ and H in Eq. (7.6) no longer holds in general. Despite our numerics for the Heisenberg interactions, it remains open whether significant destructive error interference can be achieved for higher-order product formulas in simulating other classes of Hamiltonians. In addition, although our main focus in this chapter is on real-

time simulation, it would be interesting to consider the implications of our bound for the error of the product formula in simulating imaginary time evolution, which is relevant for path integral Quantum Monte Carlo algorithms [159].

We also note that while our analysis requires $r\|\delta\| < 1/2$, our numerical calculation [see Fig. 7.1] shows that our error bound agrees well with the empirical scaling even at large values of t, where $r\|\delta\| \gg 1/2$. Therefore, we conjecture that the error bound in Eq. (7.17) is valid regardless of whether εt is less than one. If the conjecture holds, Eq. (7.19) implies that we should choose $r \propto nt/\varepsilon$ and $r \propto \sqrt{nt^3/\varepsilon}$ for $\varepsilon t \leq n$ and $\varepsilon t > n$, respectively (in the limit of large n and t). The former choice yields the same gate count $\mathcal{O}\left(n^2 t/\varepsilon\right)$ as in Eq. (7.20), but the latter choice leads to a gate count of $\mathcal{O}\left(nr\right) = \mathcal{O}\left(\sqrt{n^3 t^3/\varepsilon}\right)$, which is tighter than the estimate in Eq. (7.21). Thus, the conjecture would imply that PF1 performs as well as PF2—whose gate count is also $\mathcal{O}\left(\sqrt{n^3 t^3/\varepsilon}\right)$ [4]—in the large-time limit. We consider proving the conjecture a very interesting future direction.

Chapter 8: Faster Digital Quantum Simulation by Symmetry Protection

In this chapter, we propose an approach, using the symmetries of a target Hamiltonian, to protect its simulated dynamics against simulation errors. Given a simulation algorithm that decomposes the dynamics of the system into many small time steps (e.g., Trotterization), we interweave the simulation with unitary transformations generated by the symmetries of the system (Fig. 8.1). While these additional unitary transformations increase the gate complexity of the simulation, the error of the simulation can sometimes be reduced by several orders of magnitude, ultimately resulting in a faster quantum simulation. In addition, depending on the symmetries, the unitary transformations may be implemented using only single-qubit gates, which are considered inexpensive for implementations on near-term quantum computers.

The symmetry protection technique considered in this chapter is general and potentially applies to any algorithms that simulate the time evolution of Hamiltonians with symmetries by splitting the evolution into many time segments, including Trotterization and the higher-order product formulas [160] and more advanced algorithms such as those based on linear combinations of unitaries [8, 9, 11, 161], Lieb-Robinson bounds [21, 76], and randomized compilations [162, 162, 163]. We also provide evidence that the technique can also protect the simulation against other types of temporally correlated errors,

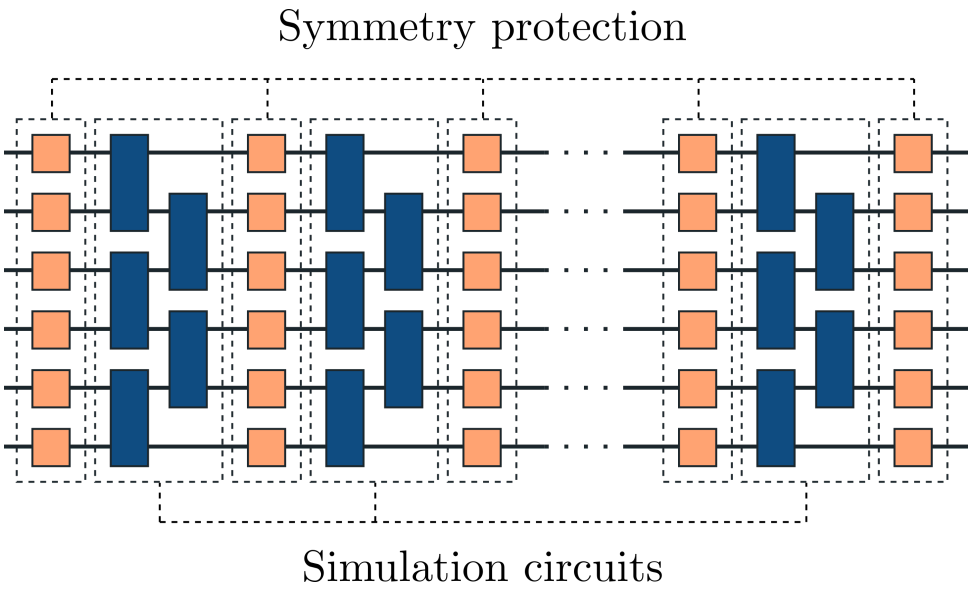

Figure 8.1: For algorithms that simulate the dynamics of quantum systems by decomposing the evolutions into many time steps, we interweave the corresponding simulation circuits (blue) with unitary transformations generated by the symmetries of the systems (orange). These transformations protect the simulations against errors that violate the symmetries, resulting in faster and more accurate simulations.

such as the $1/f$ noise commonly found in solid-state devices [164].

In addition, we draw a connection between the symmetry protection technique and the quantum Zeno effect [165, 166, 167, 168, 169, 170, 171]. In particular, the symmetry transformations, when chosen as powers of a unitary, approximately project the error of simulation into the so-called quantum Zeno subspaces, defined by the eigensubspaces of the unitary. We prove a bound on the accuracy of this approximation, exponentially improving a recent result of Ref. [171].

The structure of the chapter is as follows. In Sec. 8.1, we introduce the general technique and provides intuition for the source of error reduction. In Sec. 8.2, we derive a bound on the error of Trotterization under symmetry protection. In Sec. 8.3, we then benchmark the technique in simulating the dynamics of systems with the Heisen-

berg interactions, including the XXZ Heisenberg model with local disorder that displays a transition between thermalized and many-body localized phases, and in simulating the Schwinger model in the context of lattice field theories. In particular, we show that interweaving the simulation with random gauge transformations can significantly reduce the probability of a state leaking to outside the physical subspace due to the simulation error, extending the results of Ref. [172] to digital quantum simulation. We then demonstrate in Sec. 8.4 how the technique may protect the simulation against other types of coherent, temporally correlated errors, such as the low-frequency noise typically found in experiments. Finally, we discuss several open questions in Sec. 8.5.

8.1 General Framework

We consider the task of simulating the time dynamics of a system under a time-independent Hamiltonian H. Let $U_t \equiv \exp(-iHt)$ denote the evolution unitary generated by H for time t. The symmetry protection technique applies to algorithms that simulate U_t by first dividing the evolution into many time steps (also known as *Trotter steps*), and approximate the evolution within each time step by a series of quantum gates. Examples of such algorithms include most modern quantum simulation algorithms from the Suzuki-Trotter product formulas [160] to algorithms based on linear combinations of unitaries [8, 9, 11, 161]. In this chapter, we focus our theoretical analysis on the first-order Trotterization algorithm for simplicity (Sec. 8.2) and benchmark the performance of symmetry protection on other algorithms numerically (Sec. 8.3.2). To be more precise, let r denote the number of steps and $\delta t = t/r$ denote the length of each time step. These

algorithms then simulate $U_{\delta t}$ by a series of elementary quantum gates $S_{\delta t}$, i.e.

$$U_t = U_{\delta t}^r \approx S_{\delta t}^r. \tag{8.1}$$

The approximation of $U_{\delta t}$ by $S_{\delta t}$ introduces an error that is small for small δt. However, errors typically accumulate after many Trotter steps, resulting in a total additive error $\|U_t - S_{\delta t}^r\|$ that, in the worse case, scales linearly with the number of Trotter steps r at fixed δt. Equivalently, for a fixed total time t, to reduce the total error, we would have to decrease the Trotter step size δt, effectively increasing the number of Trotter steps r, and thus require more elementary quantum gates to run the simulation.

We refer to the simulation in Eq. (8.1) as the *raw* simulation. By exploiting symmetries of the system, we will see that we can substantially reduce the total error ε of the simulation without significantly increasing the gate count, ultimately resulting in faster quantum simulation for the same total error budget. For that, we assume that the Hamiltonian is invariant under a group of unitary transformations, which we denote by $\mathcal{S}$. Explicitly, we assume that

$$[C, H] = 0 \quad \forall\, C \in \mathcal{S}. \tag{8.2}$$

The group $\mathcal{S}$ represents a symmetry of the system. Instead of simply approximating $U_{\delta t}$ by the circuit $S_{\delta t}$, we "rotate" each implementation of $S_{\delta t}$ by a *symmetry transformation*

$C_k \in \mathcal{S}$ $(i = k, \ldots, r)$ so that the approximation in Eq. (8.1) now reads

$$U_t \approx \prod_{k=1}^{r} C_k^\dagger S_{\delta t} C_k. \tag{8.3}$$

We refer to Eq. (8.3) as a *symmetry-protected* (SP) simulation. The right-hand side in Eq. (8.3) represents a circuit that, at first, looks more expensive than Eq. (8.1) due to the additional implementation of the transformations C_k. However, for the same r, the total error in Eq. (8.3) could be much smaller than the Eq. (8.1). Effectively, to meet the same error tolerance, Eq. (8.3) may require a much smaller number of steps r, and hence fewer implementations of $S_{\delta t}$, than the raw approximation in Eq. (8.1). Moreover, because many symmetries—the gauge symmetries in lattice field theories for example—are spatially local, each C_k only involves a small number of nearest-neighboring qubits and can be implemented easily in most architectures of quantum computers. Other symmetries, such as the one responsible for the conservation of the total magnetization in the Heisenberg model, are global but may be implemented as a product of only single-qubit gates, which are usually much "cheaper" to perform in experiments than their multi-qubit counterparts.

In the remainder of this section, we provide some intuition, using lowest-order arguments, for the error reduction in simulations under symmetry protection. We later derive rigorous error bounds in Sec. 8.2.

8.1.1 Lowest-Order Arguments

To build an intuition for the symmetry protection, we consider the effective Hamiltonian of the simulation. The aim of digital quantum simulation is to simulate the time

evolution e^{-iHt} of a Hamiltonian H. Assuming that the simulation errors are coherent, we may end up with the time evolution of a different Hamiltonian, say H_{eff}, that may be close but not the same as the targeted Hamiltonian H:

$$e^{-iHt} \xrightarrow{\text{errors}} e^{-iH_{\text{eff}}t} = e^{-i(H+V)t}, \tag{8.4}$$

where

$$V \equiv H_{\text{eff}} - H \tag{8.5}$$

quantifies the difference between the effective and the desired Hamiltonians [77]. We note that the effective Hamiltonian H_{eff} typically depends on the time step δt [See Lemma 10].

With $S_{\delta t} = \exp(-iH_{\text{eff}}\delta t)$ in Eq. (8.3), we can rewrite the simulation as

$$\begin{aligned}\prod_k^r C_k^\dagger S_{\delta t} C_k &= \prod_{k=1}^r e^{-iC_k^\dagger H_{\text{eff}} C_k \delta t} \\ &= \prod_{k=1}^r e^{-i(H + C_k^\dagger V C_k)\delta t},\end{aligned} \tag{8.6}$$

where we have used the unitarity of C_k to move the unitaries to the exponents and exploited the commutativity $[C_k, H] = 0$ from our assumption to simplify the expression. Assuming that the error $\|V\|$ is small, we can use the Baker-Campbell-Hausdorff (BCH) formula to combine the exponents in Eq. (8.6) (to the leading order):

$$\prod_{k=1}^r e^{-i(H + C_k^\dagger V C_k)\delta t} \approx e^{-i\left(H + \frac{1}{r}\sum_{k=1}^r C_k^\dagger V C_k\right)t} = e^{-i\overline{H}_{\text{eff}}t}. \tag{8.7}$$

Compared to the desired evolution e^{-iHt}, we can identify the error of the entire simulation (ignoring the error from the BCH approximation for now) as

$$\bar{V} \equiv \frac{1}{r} \sum_{k=1}^{r} C_k^\dagger V C_k. \tag{8.8}$$

Roughly speaking, the error of the entire simulation, given by Eq. (8.8), can be interpreted as the average of the error in each step of the simulation. To illustrate the effect of the symmetry protection, we could imagine V as a vector in the space of operators and $C_k^\dagger V C_k$ is a version of the vector rotated around an axis specified by C_k. The total error is then analogous to a walker that, in each step, walks a distance $\|V\|$ in the space of operators towards the direction corresponding to C_k (Fig. 8.2).

Without the symmetry protection (i.e. $C_k = \mathbb{I}$ for all k), the walker keeps walking in the same direction and its total distance after r steps scales as $\mathcal{O}(r)$, resulting in the averaged error $\|\bar{V}\|$ of the same order as $\|V\|$. On the other hand, under the symmetry protection, the walker walks in a possibly different direction in each step, resulting in a smaller total distance (and thus a smaller averaged error.)

In particular, if the walker in each step walks towards a uniformly random direction in the space of operators (which is sometimes the result of choosing C_k at random), its total distance should only scale as $\mathcal{O}(\sqrt{r}\|V\|)$ after r steps. The averaged error $\|\bar{V}\|$ would then scale as $\mathcal{O}(\|V\|/\sqrt{r})$, decreasing with the number of steps of the simulation. Additionally, if we could design a set of optimal symmetry transformations that makes the walker return to the origin after a fixed number of steps, we would end up with a total distance that does not increase with r and an averaged error $\|\bar{V}\|$ that decreases with r as

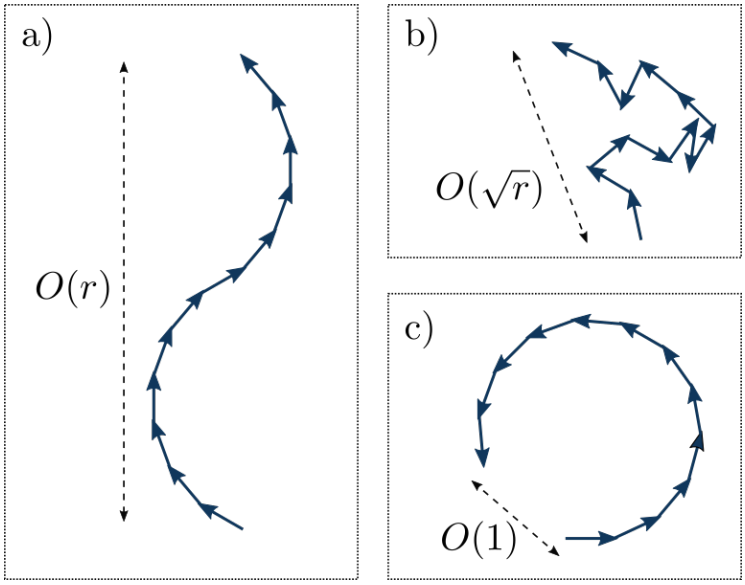

Figure 8.2: The total error of the simulation is analogous to the average distance a walker walks in r steps of the simulation. In each time step, the walker walks a small distance along a vector representing the error operator in the space of operators. a) Without any symmetry protection, the walker keeps walking towards almost the same direction, resulting in a total distance that scales linearly with the number of steps r, corresponding to the total error scaling as $\mathcal{O}(1)$. b) The symmetry transformations make the walker walk in a possibly different direction in every time step. When the direction is uniformly random (see Sec. 8.3.1.1 and Fig. 8.3 for an example), the total distance only scales as $\mathcal{O}(\sqrt{r})$, resulting in the total error scaling as $\mathcal{O}(1/\sqrt{r})$. c) Sometimes, it is possible to design an optimal set of symmetry transformations that makes the walker return to the origin [See Eq. (8.38) for an example], resulting in an $\mathcal{O}(1/r)$ error for the entire simulation.

$\mathcal{O}(\|V\|/r)$. We derive rigorous bounds to support this intuition in Sec. 8.2.

The aim of the symmetry protection technique is to choose the symmetry transformations C_k that minimize the error in Eq. (8.8). While each C_k may be chosen independently of the others, we will sometimes focus our attention on a special construction that requires $C_k = C_0^k$ for some $C_0 \in \mathcal{S}$. This choice for the transformations result in a simpler simulation circuit, i.e.

$$U_t \approx C_0^{\dagger r}(S_{\delta t}C_0)^r, \tag{8.9}$$

which corresponds to applying the same symmetry transformation C_0 alternatively with the implementations of the simulating circuit $S_{\delta t}$, followed by a final application of $C_0^{\dagger r}$ to negate the effect of C_0 on the correct evolution. We could either draw C_0 randomly from the symmetry group $\mathcal{S}$ or infer an optimal choice of C_0 from the structure of the error V [See Eq. (8.38) for an example]. We analyze the error bounds for the simulation under the protection from this special construction in Sec. 8.2 and present similar analysis for the general scenario in Appendix F.3.

It is worth noting that the symmetry transformation C_0 introduced above is also analogous to the fast pulses (or "kicks") commonly used in quantum control to confine the dynamics of quantum systems [165, 166, 167, 168, 169, 170, 171]. In fact, we also show in Appendices F.1 and F.2 that a restricted version of the symmetry protection technique is exactly equivalent to frequently applying fast pulses to the systems, resulting in the error being approximately projected onto the so-called quantum Zeno subspaces. We prove a bound on the error of this approximation, exponentially improving a recent result of Ref. [171]. This quantum Zeno framework provides an alternative explanation for how quantum simulation can be improved by symmetry protection.

8.2 Faster Trotterization by Symmetry Protection

In this section, we analyze the effect of the symmetry protection on the total error of the first-order Trotterization algorithm. Suppose the Hamiltonian $H = \sum_{\mu=1}^{L} H_\mu$ is a sum of L Hamiltonian terms H_μ such that each $e^{-iH_\mu \delta t}$ can be readily simulated on quantum

computers. For readability, we define the following quantities

$$\alpha \equiv \sum_{\mu=1}^{L} \sum_{\nu=\mu+1}^{L} \|[H_\nu, H_\mu]\|, \tag{8.10}$$

$$\beta \equiv \sum_{\mu=1}^{L} \sum_{\nu=\mu+1}^{L} \sum_{\nu'=\nu}^{L} \|[H_{\nu'}, [H_\nu, H_\mu]]\|, \tag{8.11}$$

that depend only on the commutators between the terms of the Hamiltonian. We will also use the standard Bachmann-Landau big-O and big-Θ notations in analyzing the asymptotic scalings of the errors with respect to n, t, and r. For reference, $\alpha = \mathcal{O}(n)$ and $\beta = \mathcal{O}(n)$ in a system of n nearest-neighbor interacting particles [77].

Given a set of symmetry transformations $\mathcal{C} = \{C_k : k = 1, \ldots, r\}$, we define

$$\bar{A} \equiv \frac{1}{r} \sum_{k=1}^{r} C_k^\dagger A C_k \tag{8.12}$$

as the version of an operator A averaged over the rotations induced by C_k.

The first-order Trotterization algorithm approximates $\exp(-iH\delta t)$ by

$$S_{\delta t} = \prod_{\mu=1}^{L} e^{-iH_\mu \delta t}, \tag{8.13}$$

where $\prod_{\mu=1}^{L} U_\mu \equiv U_L \ldots U_2 U_1$ is an ordered product. We define H_{eff} as the generator of $S_{\delta t}$, i.e. $S_{\delta t} = \exp(-iH_{\text{eff}}\delta t)$. We prove the following lemma, providing the existence and the structure of the generator H_{eff}.

Lemma 10. *For all δt such that $\beta\delta t \leq \alpha$, $2\alpha\delta t \leq \|H\|$, and $8\delta t\|H\| \leq 1$, there exists a*

generator H_{eff} for $S_{\delta t}$ and

$$H_{eff} = H - \frac{i}{2} v_0 \delta t + \mathcal{V}(\delta t), \tag{8.14}$$

where

$$v_0 \equiv \sum_{\mu=1}^{L} \sum_{\nu=\mu+1}^{L} [H_\nu, H_\mu], \tag{8.15}$$

$\mathcal{V}(\delta t)$ is an operator satisfying $\|\mathcal{V}(\delta t)\| \leq \chi \delta t^2$ and

$$\chi \equiv \beta + 32\alpha \|H\|. \tag{8.16}$$

We provide the proof of Lemma 10 in Appendix F.4. The essence of Lemma 10 is that the error of the simulation, defined as $V \equiv H_{\text{eff}} - H$, is given by

$$V = -\frac{i}{2} v_0 \delta t + \mathcal{O}\left(\chi \delta t^2\right), \tag{8.17}$$

and it follows that $\|V\| \leq \frac{1}{2}\alpha \delta t + \chi \delta t^2$.

We now consider the effect of protecting the simulation with a set of symmetry transformations $\{C_k : k = 1, \ldots, r\}$. Under this symmetry protection, each circuit $S_{\delta t}$ is replaced by

$$S_{\delta t} \to C_k^\dagger S_{\delta t} C_k = e^{-iC_k^\dagger H_{\text{eff}} C_k \delta t} = e^{-i(H + C_k^\dagger V C_k)\delta t}, \tag{8.18}$$

where we have used $[C_k, H] = 0$ to simplify the expression. The full simulation becomes

$$\prod_{k=1}^{r} C_k^\dagger S_{\delta t} C_k = \prod_{k=1}^{r} e^{-i(H + C_k^\dagger V C_k)\delta t}. \tag{8.19}$$

In the following analysis, we further assume that the symmetry transformations C_k have the form $C_k = C_0^k$, where C_0 is a symmetry transformation drawn from the symmetry group $\mathcal{S}$ (We extend these results to general symmetry transformations in Appendix F.3.) Let $\{e^{-i\phi_\mu} : 1 \le \mu \le m\}$ denote the distinct eigenvalues of C_0 and

$$\overline{H}_{\text{eff}} = H + \frac{1}{r}\sum_{k=1}^{r} C_k^\dagger V C_k = H + \overline{V}. \tag{8.20}$$

Lemma 11. *If $m \ge 2$, we have*

$$\begin{aligned} &\|\prod_{k=1}^{r} C_k^\dagger e^{-iH_{\text{eff}}\delta t} C_k - e^{-i\overline{H}_{\text{eff}} t}\| \\ &\le \frac{2\xi\sqrt{m}(\|H\| + \|V\|)\|V\| t^2 \log r}{r}, \end{aligned} \tag{8.21}$$

where

$$\xi \equiv \max_{\mu \ne \nu} |\sin\left(\frac{\phi_\mu - \phi_\nu}{2}\right)|^{-1} \tag{8.22}$$

is the inverse spectral gap *that depends on the eigenvalues of C_0.*

The proof of Lemma 11 follows from Lemma 37 in Appendix F.2. We note that the bound in Lemma 11 depends on m, the number of unique eigenvalues of C_0, which could

be a constant, e.g. when C_0 is generated by local symmetries, or depend on the system size, e.g. when C_0 corresponds to generic rotations generated by global symmetries. We also note that the inverse spectral gap ξ could be large if C_0 is nearly degenerate and one should take this effect into account when choosing the unitary C_0.

Lemma 11 says that, up to the error given in Eq. (8.21), the simulation under the symmetry protection is effectively described by $\overline{H}_{\text{eff}}$. In particular, the total error of the Hamiltonian under the symmetry protection is

$$\overline{V} = \overline{H}_{\text{eff}} - H = \frac{1}{r}\sum_{k=1}^{r} C_k^\dagger V C_k \tag{8.23}$$

$$= \frac{-i}{2}\underbrace{\frac{1}{r}\sum_{k=1}^{r} C_k^\dagger v_0 C_k}_{=\overline{v_0}}\,\delta t + \underbrace{\frac{1}{r}\sum_{k=1}^{r} C_k^\dagger \mathcal{V} C_k}_{=\bar{\mathcal{V}}}, \tag{8.24}$$

where we have replaced the expression of V from Lemma 10. Note that $\|\bar{\mathcal{V}}\| \le \|\mathcal{V}\|$ by the triangle inequality. Using the identity

$$\|e^{-i\overline{H}_{\text{eff}}t} - e^{-iHt}\| \le \|\overline{H}_{\text{eff}} - H\|t = \|\overline{V}\|t, \tag{8.25}$$

we arrive at the following bound on the total error of the simulation.

Theorem 10 (Quantum simulation by symmetry protection)**.** *Assuming that* $\beta\delta t \le \alpha$, $2\alpha\delta t \le \|H\|$, *and* $8\delta t\|H\| \le 1$, *the total error of simulation under the symmetry protec-*

tion from $\{C_k = C_0^k : C_0 \in \mathcal{S}, k = 1, \ldots, r\}$ *can be bounded as*

$$\begin{aligned} \varepsilon &\equiv \|\prod_{k=1}^{r} C_k^\dagger S_{\delta t} C_k - e^{-iHt}\| \\ &\leq \|\overline{v_0}\| \frac{t^2}{2r} + \chi \frac{t^3}{r^2} + \kappa \frac{t^3 \log r}{r^2}, \end{aligned} \tag{8.26}$$

where

$$\chi \equiv \beta + 32\alpha\|H\|, \quad \kappa \equiv 48\xi\sqrt{m}\alpha\|H\|, \tag{8.27}$$

m *is the number of distinct eigenvalues of* C_0*, and* ξ *is the inverse spectral gap defined in Eq.* (8.22).

The proof of Theorem 10 follows immediately from Lemma 11 and Eq. (8.25) [See Appendix F.5 for the detailed calculations]. The key feature of Theorem 10 is that, to the lowest-order in $\frac{t}{r}$, the error scales with $\|\overline{v_0}\|$ instead of $\|v_0\|$. Since

$$\|\overline{v_0}\| = \|\frac{1}{r}\sum_{k=1}^{r} C_k^\dagger v_0 C_k\| \tag{8.28}$$

is generally smaller than $\|v_0\|$ when $[C_k, v_0] \neq 0$, we expect a smaller simulation error under the symmetry protection.

For demonstration, we consider the simulation of a Hamiltonian H that is a sum of nearest-neighbor interactions on n particles. It is straightforward to verify that for this Hamiltonian, $\|H\| = \mathcal{O}(n)$, $\|v_0\| \leq \alpha = \mathcal{O}(n)$, $\beta = \mathcal{O}(n)$, and $\chi = \mathcal{O}(n^2)$. We will also assume that the number of distinct eigenvalues of the C_0 is $m = \mathcal{O}(1)$ (corresponding

to local symmetries or highly degenerate transformations) which results in $\kappa = \mathcal{O}(n^2)$. We will estimate the required number of steps r—a good proxy for the gate count [1]—for simulations with and without the symmetry protection.

The first scenario corresponds to an unprotected simulation, where $\overline{v_0} = v_0$. The total error then scales as

$$\varepsilon = \mathcal{O}\left(\frac{nt^2}{r}\right) + \mathcal{O}\left(\frac{n^2 t^3 \log r}{r^2}\right). \tag{8.29}$$

To meet a fixed error tolerance ε, we would have to choose the number of steps $r = \Theta(nt^2/\varepsilon)$.

On the other hand, with symmetry protection, we later show that it is sometimes possible to make $\overline{v_0}$ vanish completely, making the higher order terms the dominant contribution to the total error [See Eq. (8.38) for an example]. For nearest-neighbor interactions, the total error is now

$$\varepsilon = \mathcal{O}\left(\frac{n^2 t^3 \log r}{r^2}\right), \tag{8.30}$$

which decreases quadratically with r. As a result, we only need

$$r = \tilde{\Theta}\left(\frac{nt^{3/2}}{\sqrt{\varepsilon}}\right), \tag{8.31}$$

where $\tilde{\Theta}(\cdot)$ is $\Theta(\cdot)$ up to a logarithmic correction. Note that this choice of r also satisfies

[1] In most quantum computer architectures, implementing one Trotter step costs the same amount of gates regardless of the time step δt. Here, we also assume that the cost of implementing each symmetry transformation is negligible compared to the simulation circuit. Therefore, given a fixed n, the Trotter number r is proportional to the gate count of the simulation.

the conditions in Theorem 10 when $t/\varepsilon > 1$. Compared to the unprotected simulation, the symmetry protection results in a factor of $\sqrt{t/\varepsilon}$ improvement in the required number of steps. At $\varepsilon = 0.01$, the improvement in the scaling with ε alone would result in about a factor of ten reduction in the gate count of the simulation.

Finally, we consider a scenario where $\|\overline{v_0}\| \propto \|v_0\|/r^\gamma$ for some $\gamma \in (0, 1)$. We provide an example of such a scaling in Sec. 8.3.1.1, where drawing the unitary transformations C_k randomly from the symmetry group results in a scaling with $\gamma = 0.5$. This scaling of $\|\overline{v_0}\|$ results in the total error

$$\varepsilon = \mathcal{O}\left(\frac{nt^2}{r^{1+\gamma}}\right) + \mathcal{O}\left(\frac{n^2t^3 \log r}{r^2}\right). \tag{8.32}$$

Hence, we require

$$r = \max\left\{\Theta\left(\left(\frac{nt^2}{\varepsilon}\right)^{\frac{1}{1+\gamma}}\right), \tilde{\Theta}\left(\frac{nt^{3/2}}{\sqrt{\varepsilon}}\right)\right\},$$

which is again better than the unprotected simulation by a factor of $\min\{(nt^2/\varepsilon)^{\gamma/(1+\gamma)}, \sqrt{t/\varepsilon}\}$.

We recall that in deriving Theorem 10, we have assumed that the symmetry transformations have the form $C_k = C_0^k$ for some C_0. We derive in Appendix F.3 a different bound for the general case where each C_k may be chosen independently. This general bound, while appearing more complicated, holds the same key feature to the bound in Theorem 10: the total error, to the lowest-order, scales with an averaged version of v_0 (under the symmetry transformations) instead of scaling with $\|v_0\|$.

8.3 Applications

In this section, we apply the symmetry protection technique to the simulation of the Heisenberg model (Sec. 8.3.1) and lattice field theories (Sec. 8.3.2). In both cases, we show that the symmetry protection results in a significant error reduction and thereby gives faster quantum simulation.

In particular, we use the simulation of the homogeneous Heisenberg model in Sec. 8.3.1.1 to demonstrate the improvement on the total error scaling as a function of the number of steps r when the simulation is protected by a random set of unitary transformations and by an optimally chosen set. In Sec. 8.3.1.2, we estimate the required number of Trotter steps as a proxy for the gate count in simulating an instance of the Heisenberg model, commonly found in the studies of the many-body localization phenomenon. Finally, in Sec. 8.3.2, we consider the probability of the state leaking to unphysical subspaces in the digital simulation of the Schwinger model and show that the symmetry protection from the local gauge symmetries can suppress this leakage by a few orders of magnitude.

8.3.1 Heisenberg Interactions

In this section, we use the symmetries in the Heisenberg model to protect its simulation using the first-order Trotterization. A Heisenberg model of n spins can be described

by the Hamiltonian

$$H = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(J_{ij}^{(x)} X_i X_j + J_{ij}^{(y)} Y_i Y_j + J_{ij}^{(z)} Z_i Z_j \right) + \sum_{i=1}^{n} h_i Z_i, \tag{8.33}$$

where X_i, Y_i, Z_i are the Pauli matrices acting on site i, $J_{ij}^{(x,y,z)}$ represent the interaction strengths between the spins, and h_i correspond to the strengths at site i of an external magnetic field pointing in the z direction. The Heisenberg model provides a good description for the behavior of magnetic materials in the presence of external magnetic fields. Depending on several factors, including the signs of the interactions and the dimensions of the system, the Heisenberg model may undergo a quantum phase transition as we increase the strength of the external magnetic field. Several important instances of the Heisenberg model includes the homogeneous Heisenberg model ($J^{(x)} = J^{(y)} = J^{(z)}$), the XXZ model ($J^{(x)} = J^{(y)}$) with local disorder, and the Ising model ($J^{(y)} = J^{(z)} = 0$). In the following subsections, we will consider two pedagogical instances of Eq. (8.33) with SU(2) and U(1) symmetries respectively and demonstrate how the symmetry protection helps reduce the error in simulating the dynamics of these systems even as they move across critical points.

8.3.1.1 Homogeneous, Random Heisenberg Interactions

We first consider a pedagogical toy model where interactions in Eq. (8.33) are homogeneous, i.e. $J_{ij}^{(x)} = J_{ij}^{(y)} = J_{ij}^{(z)} = J_{ij}$ for all $1 \leq i < j \leq n$, but each J_{ij} is chosen

independently at random between $[-1, 1]$. In addition, we assume that $h_i = 0\ \forall i$, i.e. there is no external magnetic field. In this case, Eq. (8.33) simplifies to

$$H = \underbrace{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} J_{ij} X_i X_j}_{\equiv H_X} + \underbrace{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} J_{ij} Y_i Y_j}_{\equiv H_Y} + \underbrace{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} J_{ij} Z_i Z_j}_{\equiv H_Z} . \tag{8.34}$$

The combination of homogeneous interactions and no external magnetic field make Eq. (8.34) invariant under $\mathcal{S} = \{W^{\otimes n} : W \in \text{SU(2)}\}$, which contains unitaries that—in the Bloch sphere—simultaneously rotate each spin by the same angle.

To simulate the evolution U_t under Eq. (8.34), we could use the first-order Trotterization to approximate

$$U_t = \left(e^{-iH\delta t}\right)^r \approx \left(e^{-iH_X\delta t} e^{-iH_Y\delta t} e^{-iH_Z\delta t}\right)^r \tag{8.35}$$

by a product of evolutions of individual terms of the Hamiltonian. The number of Trotter steps r and the time step $\delta t = t/r$ determine the error of the simulation. We refer to this approach as the *raw* Trotterization. To protect this simulation, we insert unitaries drawn from the symmetry group $\mathcal{S}$ in between the Trotter steps, resulting in the simulation

$$U_t = \left(e^{-iH\delta t}\right)^r \approx \prod_{k=1}^{r} C_k^\dagger e^{-iH_X\delta t} e^{-iH_Y\delta t} e^{-iH_Z\delta t} C_k, \tag{8.36}$$

where $\{C_1, \ldots, C_r\} \equiv \mathcal{C}$ is a subset of the symmetry group $\mathcal{S}$. Recall that the total error

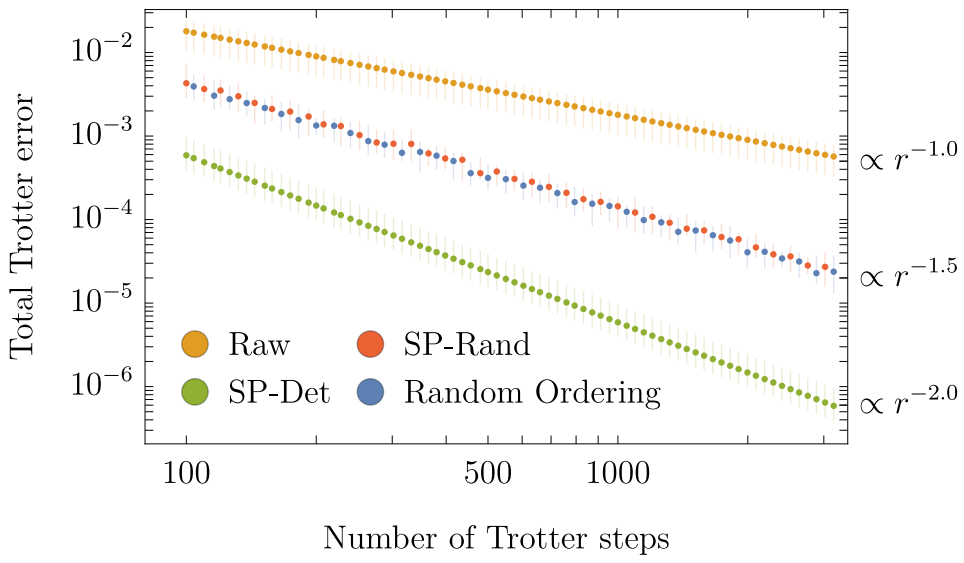

Figure 8.3: The total error in simulating the Hamiltonian Eq. (8.34) at $n = 4$ for a fixed evolution time $t = 1$ as a function of the Trotter number r using four different schemes: the raw first-order Trotterization ("Raw"), the first-order Trotterization protected by a random set symmetry transformation ("SP-Rand"), the first-order Trotterization protected by the optimal set in Eq. (8.38) ("SP-Det"), and the random-ordering scheme in Ref. [163] ("Random Ordering"). We indicate the scalings obtained from power-law fits to the right of the plot. We repeat the simulation 100 times, each with a different set of randomly generated interactions J_{ij}. The dots correspond to the median of the errors at each value of r and the bars represent the corresponding 25%-75% percentiles regions.

of this symmetry-protected simulation is given by Theorem 10, with the lowest-order error being

$$\frac{t^2}{2r}\|\overline{v_0}\| = \frac{t^2}{2r}\|\sum_{k=1}^{r} C_k^\dagger v_0 C_k\|, \tag{8.37}$$

where $v_0 = [H_Y, H_X] + [H_Z, H_X] + [H_Z, H_Y]$ comes from the leading contribution to the error in one Trotter step. Different choices of the set $\mathcal{C}$ lead to different total error of the simulation.

For minimal calculational overhead, we could choose each C_k independently and uniformly at random from S (i.e. $C_k = W_k^{\otimes n}$ where W_k is a Haar random unitary on

the single-qubit Bloch sphere.) The sum in Eq. (8.37) is then the sum of v_0, each rotated under a random unitary. This is analogous to the total error being a random walker that, in each time step, "walks" a distance $\|v_0\|$ in a random direction (See Fig. 8.2). From this analogy, we then expect $\|\overline{v_0}\| \propto \|v_0\|/\sqrt{r}$ (to the lowest-order). Therefore, we expect the total error of this scheme to decrease with r as $\mathcal{O}\left(r^{-3/2}\right)$ (at fixed total time t).

While randomly choosing the unitary transformation set $\mathcal{C}$ requires little to no knowledge about the error operator v_0, one can expect that this choice of $\mathcal{C}$ is not optimal. Indeed, by further exploiting the structure of v_0, we can construct a set of transformations $\mathcal{C}$ that makes Eq. (8.37) vanishes entirely. One such choice is $C_k = C_0^k$ for $k = 1, \ldots, r$, where

$$C_0 = U_H^{\otimes n}, \tag{8.38}$$

and U_H is the single-qubit Hadamard matrix. Alternatively, we could also write

$$C_k = \begin{cases} \mathbb{I} & \text{if } k \equiv 0 \mod 2, \\ U_H^{\otimes n} & \text{if } k \equiv 1 \mod 2, \end{cases} \tag{8.39}$$

for $k = 1, \ldots, r$. Since the Hadamard matrix switches $X \leftrightarrow Z$ and $Y \leftrightarrow -Y$, it is straightforward to verify that Eq. (8.37) vanishes for all even values of r. Therefore, the total error of the simulation is given by the next lowest order in Theorem 10, which scales with r as $\mathcal{O}\left(1/r^2\right)$.

In Fig. 8.3, we plot the total error of the simulation at $n = 4, t = 1$ as a function of

the Trotter number r for the three aforementioned scenarios: the first-order Trotterization without symmetry protection ("Raw"), with symmetry protection from a randomly chosen $\mathcal{C}$ ("SP-Rand"), and with symmetry protection from the optimal set $\mathcal{C}$ ("SP-Det"). The scalings of the errors as functions of r agree remarkably well with our above prediction. In addition, we also compute the total error using the randomized simulation scheme in Ref. [163], which decreases the Trotter error by randomizing the ordering of the Hamiltonian terms in between Trotter steps. Our numerics shows that this scheme performs similarly to the simulation protected by random symmetry transformations, which are both outperformed by the optimal symmetry protection scheme.

8.3.1.2 Many-Body Localization

The homogeneous Heisenberg interactions without external fields considered in the previous section provides a good testbed for benchmarking the symmetry protection technique. In this section, we consider a more physically relevant instance of the Heisenberg model:

$$H = \sum_{i=1}^{n} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1} + \sum_{i=1}^{n} h_i Z_i, \tag{8.40}$$

where we again assume homogeneity for the coupling strengths, but $J_{ij} = 1$ only when i, j are nearest neighbors and $J_{ij} = 0$ otherwise. We also adopt the periodic boundary condition and identify the $(n+1)$th qubit as the first qubit. In addition, we add an external magnetic field with the field strength h_i, each chosen randomly between $[-h, h]$. This model describes homogeneous Heisenberg interactions with a tunable local disor-

der strength h. At low disorder h, the system evolved under Eq. (8.40) thermalizes in the long-time limit, in agreement with the Eigenstate Thermalization Hypothesis (ETH). However, as h increases, the system transitions to a many-body localized (MBL) phase where it no longer thermalizes (See [173] for a review of the many-body localization phenomenon.)

To simulate the dynamics of H, we again divide the terms of H into groups of mutually commuting terms:

$$H = \underbrace{\sum_{i=1}^{n} X_i X_{i+1}}_{\equiv H_X} + \underbrace{\sum_{i=1}^{n} Y_i Y_{i+1}}_{\equiv H_Y} + \underbrace{\sum_{i=1}^{n} Z_i Z_{i+1} + \sum_{i=1}^{n} h_i Z_i}_{\equiv H_Z}, \tag{8.41}$$

and use the first-order Trotterization similarly to Eq. (8.35). To symmetry-protect this simulation, we note that the field term breaks the SU(2) symmetry of the Heisenberg interactions, leaving the system invariant under a U(1) symmetry only. The symmetry group $\mathcal{S} = \{[\exp(-i\phi Z)]^{\otimes n} : \phi \in [0, 2\pi)\}$ is generated by the total spin components along the z axis $S_z \equiv \sum_{i=1}^{n} Z_i$.

While selecting the unitary transformations C_k from this U(1) symmetry is no longer sufficient to completely eliminate the lowest-order error—as we have done in the previous section—we can still expect significantly reduction of the total error due to the symmetry protection and thus a lower gate count for the simulation. In Fig. 8.4, we plot the number of Trotter steps r in simulating the dynamics of Eq. (8.40) for time $t = n$ at different values of the disorder h that correspond to the ETH and the MBL phases. The required numbers of steps are computed at each n by binary searching for the minimum r

such that the total error of the simulation does not exceed $\varepsilon = 0.01$. Figure 8.4 shows that protecting the simulation with the U(1) symmetry results in several times reduction in the number of Trotter steps for all values of n. In addition, the Trotter number under symmetry protection also appears to scale better with the system size than in the raw simulation, suggesting an even greater advantage from the symmetry protection for simulating larger systems.

Out of curiosity, we study how the symmetry protection performs as the Hamiltonian moves across the ETH-MBL phase transition. In Fig. 8.5, we plot the required number of steps r in simulating the Hamiltonian of $n = 8$ qubits for time $t = n$ and error tolerance $\varepsilon = 0.01$ as we tune the Hamiltonian from the ETH to the MBL phase [174]. The improvement due to the symmetry protection appears to be unaffected by the phase transition, suggesting that the symmetry protection technique can be useful for future numerical and experimental studies of the transition.

8.3.2 Simulation of Lattice Gauge Field Theories

Quantum field theories provide another key target for quantum simulation [175]. In particular, the quantum simulation of real-time Hamiltonian dynamics, for example scattering processes [176], has attracted much attention. An important class of field theories are models with local gauge symmetry, including quantum electrodynamics, chromodynamics, and the Standard Model of particle physics in addition to many condensed matter systems. Substantial effort has gone into the study of analog [177, 178, 179] and digital [180, 181, 182, 183, 184] quantum simulation of these models.

In a gauge theory, the system is invariant under a symmetry group which acts separately at each point in space and time (see eg. [185] for a review, as well as the lattice Hamiltonian formulation, of these models). This symmetry is fundamentally a redundancy of our description of the physics which we have introduced to give a local description. The Hilbert space $\mathcal{H}$ we use to describe the system contains a subspace $\mathcal{H}_{\text{phys}}$ of the physical states, those annihilated by the gauge constraints. For example, in electrodynamics, we have the charge and gauge field degrees of freedom, and the physical states are those annihilated by the Gauss law constraint $\mathcal{G} = \nabla \cdot \mathbf{E} - \rho$, where $\mathbf{E}$ is the electric field operator and ρ is the charge density operator. There are many states in the full Hilbert space $\mathcal{H}$ which do not live in the kernel of $\mathcal{G}$, and these states are not allowed in nature. Although one can in principle work with a description strictly within the physical Hilbert space, it is in general computationally difficult to do the reduction. More importantly, this description would necessarily have a highly spatially non-local set of interactions, a major drawback in practice.

Thus in the simulation of a gauge theory we are faced with a fundamental source of possible errors: what if our dynamics takes us away from the physical Hilbert space? Although the exact Hamiltonian commutes with the gauge constraints, and thus leaves the physical space invariant, an approximate (for example, Trotterized) version of the Hamiltonian may induce leakage into the unphysical space [172, 184]. In this section, we apply the symmetry protection technique and use the gauge symmetry itself to protect the simulation against this undesirable leakage .

Explicitly, we consider the one-dimensional Schwinger model [183, 184, 186, 187, 188, 189] consisting of n sites and $n - 1$ nearest-neighbor links between the sites. We

use the formalism outlined in Ref. [184]. The Hamiltonian $H = H_0 + H_1$ consists of two terms:

$$H_0 = \sum_{i=1}^{n-1} F_i^2 - \frac{\mu}{2} \sum_{i=1}^{n} (-1)^i Z_i, \tag{8.42}$$

$$H_1 = x \sum_{i=1}^{n-1} \Bigg[\frac{1}{4}(U_i + U_i^\dagger)(X_i X_{i+1} + Y_i Y_{i+1}) + \frac{i}{4}(U_i - U_i^\dagger)(X_i Y_{i+1} - Y_i X_{i+1}) \Bigg], \tag{8.43}$$

where

$$F_i = \sum_{j=-\Lambda}^{\Lambda-1} j \left|j\right\rangle_i \left\langle j\right|_i, \tag{8.44}$$

$$U_i = \sum_{j=-\Lambda}^{\Lambda-2} \left|j+1\right\rangle_i \left\langle j\right|_i + \left|-\Lambda\right\rangle_i \left\langle \Lambda - 1\right|_i, \tag{8.45}$$

and μ, x are positive constants. Here, H_0 describes the on-site and on-link terms, H_1 describes the site-link interaction, and F_i is the electromagnetic field operator for the link that connects the ith and $(i+1)$th particles. We note that while the second term in Eq. (8.42) sometimes appears in the literature without the minus sign (see for example Ref. [183]), this discrepancy is the result of different conventions for mapping between fermions and spins and does not have any physical consequences. In a simulation, we have to put a cutoff Λ specifying the maximum excitation number for the bosonic degree of freedom on a given link.

The Hamiltonian is subjected to local symmetries generated by the gauge operators:

$$\mathcal{G}_i = F_i - F_{i-1} - Q_i, \tag{8.46}$$

where $Q_i = \frac{1}{2}\left[-Z_i + (-1)^i\right]$ counts the electric charge at site i. In particular, only states $|\psi\rangle$ that satisfy $\mathcal{G}_i = 0$ for all i are considered physical.

The physical states form a subspace $\mathcal{H}_{\text{phys}}$ which can be constructed from the kernels of the gauge operators:

$$\mathcal{H}_{\text{phys}} \equiv \bigcap\nolimits_i \text{Ker}(\mathcal{G}_i), \tag{8.47}$$

where $\text{Ker}(\mathcal{G}_i) = \{|\phi\rangle : \mathcal{G}_i |\phi\rangle = 0\}$ is the kernel of $\mathcal{G}_i$.

Due to various errors, an initially physical state may leak to unphysical subspace during the simulation. Formally, we define the leakage of a state $|\psi(t)\rangle$ at time t as

$$1 - |\langle\psi(t)| \Pi_0 |\psi(t)\rangle|, \tag{8.48}$$

where Π_0 is the projector onto the physical subspace $\mathcal{H}_{\text{phys}}$.

To simulate $e^{-iH\delta t}$ for a small time δt, we first decompose it into $e^{-iH_0\delta t}e^{-iH_1\delta t}$ using the first order Trotterization. Since both H_0, H_1 commute with $\mathcal{G}_i$, this decomposition respects the gauge symmetries and does not result in leakage from the physical subspace. However, to simulate the evolution under H_1, we need to further decompose it

into elementary quantum gates. For that, we follow the steps in Ref. [184] and write

$$U_i + U_i^\dagger = A_i + \tilde{A}_i, \tag{8.49}$$

where $A_i = \mathbb{I} \otimes \cdots \otimes \mathbb{I} \otimes X$ and $\tilde{A}_i = U_i^\dagger A_i U_i$. Similarly,

$$i(U_i - U_i^\dagger) = B_i + \tilde{B}_i, \tag{8.50}$$

where $B_i = \mathbb{I}\otimes\cdots\otimes\mathbb{I}\otimes Y$ and $\tilde{B}_i = U_i^\dagger B_i U_i$. This representation allows us to decompose the evolution

$$\begin{aligned} e^{-iH_0\delta t}e^{-iH_1\delta t} \approx S_{\delta t} \equiv e^{-iH_0\delta t} & \\ \cdot\prod_i e^{-\frac{1}{4}ix\delta t A_i X_i X_{i+1}}e^{-\frac{1}{4}ix\delta t \tilde{A}_i X_i X_{i+1}} & \\ \cdot e^{-\frac{1}{4}ix\delta t A_i Y_i Y_{i+1}}e^{-\frac{1}{4}ix\delta t \tilde{A}_i Y_i Y_{i+1}} & \\ \cdot e^{-\frac{1}{4}ix\delta t B_i X_i Y_{i+1}}e^{-\frac{1}{4}ix\delta t \tilde{B}_i X_i Y_{i+1}} & \\ \cdot e^{+\frac{1}{4}ix\delta t B_i Y_i X_{i+1}}e^{+\frac{1}{4}ix\delta t \tilde{B}_i Y_i X_{i+1}}, & \end{aligned} \tag{8.51}$$

into a product of three-qubit gates that can be readily implemented on quantum computers [184]. Note that the cost of simulating $e^{-\frac{1}{4}ixt\tilde{A}_i X_i X_{i+1}}$ is that of approximating $e^{-\frac{1}{4}ixtA_i X_i X_{i+1}}$, plus the cost of implementing U_i, $U_i^\dagger$:

$$e^{-\frac{1}{4}ix\delta t\tilde{A}_i X_i X_{i+1}} = U_i^\dagger e^{-\frac{1}{4}ix\delta t A_i X_i X_{i+1}} U_i. \tag{8.52}$$

The entire raw first-order Trotterization simulation of e^{-iHt} becomes

$$e^{-iHt} \approx S_{\delta t}^{r}. \tag{8.53}$$

Similarly to the Heisenberg model, we could protect this simulation by interweaving the Trotter steps with symmetry transformations of the system:

$$e^{-iHt} \approx \prod_{k=1}^{r} C_k^{\dagger} S_{\delta t} C_k, \tag{8.54}$$

where C_k are generated by the gauge operators in Eq. (8.46). Specifically, we choose

$$C_k = \prod_{i=1}^{n} \exp(-i\phi_{k,i}\mathcal{G}_i) \tag{8.55}$$

for some angles $\phi_{k,i}$.

Since we truncate the spectrum of each bosonic link to $[-\Lambda+1, \Lambda]$, the transformations C_k in general commute with the Hamiltonian of the system only if we choose $\phi_{k,i} = m_{k,i}\pi/\Lambda$, where $m_{k,i}$ are integers. These transformations effectively form a $\mathbb{Z}_{2\Lambda}$ symmetry of the truncated Hamiltonian [190, 191]. However, the U(1) symmetry can be recovered by assuming a vanishing background field and choosing a large enough cutoff level Λ such that, in the physical subspace, the bosonic links never "see" the cutoff. More rigorously, if $\Lambda > n/2 + 1$, the transformations C_k commute with $\Pi_0 H \Pi_0$, where Π_0 is the projection onto the physical subspace $\mathcal{H}_{\text{phys}}$, for all angles $\phi_{k,i} \in [0, 2\pi)$.

In Fig. 8.6, we plot the leakage outside the physical subspace due to the Trotter error during simulations with and without symmetry protection. Specifically, we simulate the

evolution of the ground state of the Schwinger model with 4 sites and 3 links at $x = 0.6$, $\mu = 0.1$, $\delta t = 0.01$, and $\Lambda = 4$. This choice of Λ ensures that the Hamiltonian has a $\mathbb{Z}_8$ symmetry in general and a U(1) symmetry when restricted to the physical subspace. We consider two choices of the angles $\phi_{k,i}$: $\phi_{k,i} = k\phi_{1,i}$ ("Uniform"), for some randomly chosen $\phi_{1,i}$, and $\phi_{k,i}$ chosen independently at random for each k ("Random"). We repeat the simulation 100 times, each with a different choice of the angles.

Figure 8.6 shows that the symmetry protection can reduce the leakage to the unphysical subspace by several orders of magnitude compared to a raw simulation. While the leakage builds up in a raw simulation, the uniform choice of the transformations from the $U(1)$ symmetry results in bounded leakage during the entire simulation. This feature resembles the optimal symmetry protection discussed in Sec. 8.3.1.1 for the Heisenberg models, where the symmetry protection suppresses the simulation error nearly completely. Different choices of the symmetry transformations also affect performance of the scheme differently. While the random choices of transformations from $\mathbb{Z}_8$ and U(1) have the same effect on the leakage, the uniform choice of transformations from $\mathbb{Z}_8$ performs significantly worse than the U(1) counterpart. This discrepancy is likely because we have only eight choices for the $\mathbb{Z}_8$ symmetry transformations, whereas with the U(1) symmetry the number of choices is theoretically infinite. Effectively, the symmetry group $\mathbb{Z}_8$ has less freedom and, therefore, is less effective in averaging out the simulation error than U(1).

While our analysis in Sec. 8.2 focuses on the application to the first-order Trotterization algorithm, it is clear from the analysis that the symmetry protection will suppress any simulation errors that violate the symmetries of the system, including errors

from more advanced algorithms. To support this claim, we provide in Fig. 8.7 numerical evidence of the symmetry protection suppressing the leakage to unphysical subspace in simulating the Schwinger model using the second-order Suzuki-Trotter formula, the fourth-order Suzuki-Trotter formula [160], and a multi-product formula implemented via a linear combination of unitaries [161].

Given a Hamiltonian $H = \sum_{\nu=1}^{L} H_\nu$ being a sum of L terms, the second-order Suzuki-Trotter formula simulate the time evolution $e^{-iH\delta t}$ by

$$P_2(\delta t) = \prod_{\nu=1}^{L} e^{-iH_\nu \frac{\delta t}{2}} \cdot \prod_{\nu=L}^{1} e^{-iH_\nu \frac{\delta t}{2}}, \tag{8.56}$$

which is correct up to an $\mathcal{O}\left(\delta t^3\right)$ error. The formula can be generalized to any even order $p \geq 2$ through a recursive construction [160]:

$$P_p(\delta t) = P_{p-2}(\kappa_p \delta t)^2 P_{p-2}((1-4\kappa_p)\delta t) P_{p-2}(\kappa_p \delta t)^2, \tag{8.57}$$

where $\kappa_p = 1/(4 - 4^{1/p})$. The pth-order formula approximates $e^{-iH\delta t}$ up to an error $\mathcal{O}\left(\delta t^{p+1}\right)$. Given a small δt, the formulas can be made arbitrarily accurate by increasing p at the cost of increasing the gate count exponentially with p.

In contrast, multi-product formulas [192] enable the construction of any pth-order approximations using only poly(p) quantum gates by approximating the time evolution by sums of product formulas. Asymptotically, the gate counts of the multi-product formulas have polylogarithmic dependence on the inverse of the error tolerance. Therefore, when used as a subroutine in the Lieb-Robinson-bound-based algorithm [76], the multi-product

formulas also result in asymptotically optimal gate counts, up to polylogarithmic corrections, in simulating geometrically local systems. Because a sum of product formulas is generally not unitary, it must be implemented using techniques such as linear combinations of unitaries (LCU) [161], which encodes the multi-product formula into a unitary acting in a larger Hilbert space. Here, we will simulate the Schwinger model using a multi-product formula constructed by Childs and Wiebe [161]:

$$M(\delta t) = \frac{16}{15} P_2(\delta t/4)^4 - \frac{1}{15} P_2(\delta t), \tag{8.58}$$

which is a linear combination of two second-order product formulas.

Figure 8.7 plots the leakage to the unphysical subspace during the simulation at $n = 4, x = 0.6, \mu = 0.1, \delta t = 0.01$, and $\Lambda = 4$ using the second-order Suzuki-Trotter formula, the fourth-order Suzuki-Trotter formula, and the multi-product formula [Eq. (8.58)] with and and without symmetry protection. We implement the multi-product formula using LCU and an additional ancillary qubit. For the considered algorithms, the numerics show similar features to Fig. 8.6, where the symmetry protection suppresses the leakage by several orders of magnitude and, in particular, the uniform choice of transformations results in bounded errors throughout the simulation. The figure therefore demonstrates the generality of our approach in protecting digital quantum simulations against errors that violate symmetries of the target system. We note that the dips in the leakage of the raw simulations are likely due to the small system size considered in the simulations.

8.4 Additional Protection Against Experimental Errors

So far, we have demonstrated that symmetries in quantum systems can be used to suppress the simulation error of the Trotterization algorithm. In this section, we discuss how the technique may also protect the simulation against other types of error, including the experimental errors that may arise in the implementation of Trotterization.

In our earlier derivation, we show that the lowest-order contribution to the total error is

$$\|\bar{v}_0\| = \frac{1}{r}\|\sum_{k=1}^{r} C_k^\dagger v_0 C_k\|, \tag{8.59}$$

where v_0 is the lowest-order error from the simulation algorithm. This derivation applies equally well for the case when the error v_0 comes from sources other than the approximations in the simulation algorithms.

However, in our analysis, we require that v_0 remains the same for different steps of the simulation. In other words, the error v_0 for different Trotter steps are correlated in time. In particular, an error with temporal correlation lengths being longer than the time step δt would enable us to choose the symmetry transformations such that the errors from several consecutive steps interfere destructively. Therefore, we expect the symmetry protection technique to help reduce low-frequency noises, such as the $1/f$ noise typically found in solid-state qubit systems.

We provide numerical evidence for this argument by adding temporally correlated errors to the simulation of the Schwinger model. Specifically, after each step k of the

simulation, we apply single-qubit rotations $\exp(-i\eta\ \vec{\sigma}\cdot\hat{n}_k)$ on the system, where $\eta = 0.01$ is a small angle, around a random axis $\hat{n}_k$. These rotations mimic the effect of a depolarizing channel and violate the gauge symmetries, resulting in the state leaking to the unphysical subspace. To impart temporal correlations into this noise model, we choose the random unit vectors $\hat{n}_k$ again only after every λ consecutive Trotter steps. The parameter λ therefore plays the role of the correlation length of the noise.

In Fig. 8.8, we plot the probability that the state leaks to unphysical subspace (due to the simulation error) as a function of time for several values of the correlation length λ. To study the effect of the symmetry protection technique on the added experimental noise, we use the fourth-order Trotterization in the simulation to suppress the algorithm error, making the added noise the main contributor to the leakage observed in Fig. 8.8. As expected, at $\lambda = 1$, the experimental error varies too fast between Trotter steps and is immune against the symmetry protection technique. However, the technique begins to suppress the experimental error as soon as the noise becomes temporally correlated ($\lambda > 1$) and becomes more effective as the correlation length λ increases. Even at $\lambda = 4$, we have managed to reduce error by about an of magnitude.

8.5 Discussion

In this chapter, we propose a general technique to suppress the error of quantum simulation using the symmetries available in quantum systems, ultimately resulting in faster digital quantum simulation. We have analyzed the technique when applied to the Trotterization algorithm and derived bounds on the total error of the simulation under

symmetry protection. The bound provides insights for choosing the set of unitary transformations that optimally suppress the simulation error. We then benchmarked the technique in simulating the Heisenberg model and lattice field theories. Both examples showed that the symmetry protection results in significant reduction in the total error, and thus the gate count, of the simulation. Finally, we argue that the technique can also protect digital quantum simulation against temporally correlated noise in experiments.

Lastly, we would like to note that, although our analysis focuses on digital quantum simulation, we expect the symmetry protection technique to apply equally well for analog quantum simulation and classical simulation of the dynamics of quantum systems.

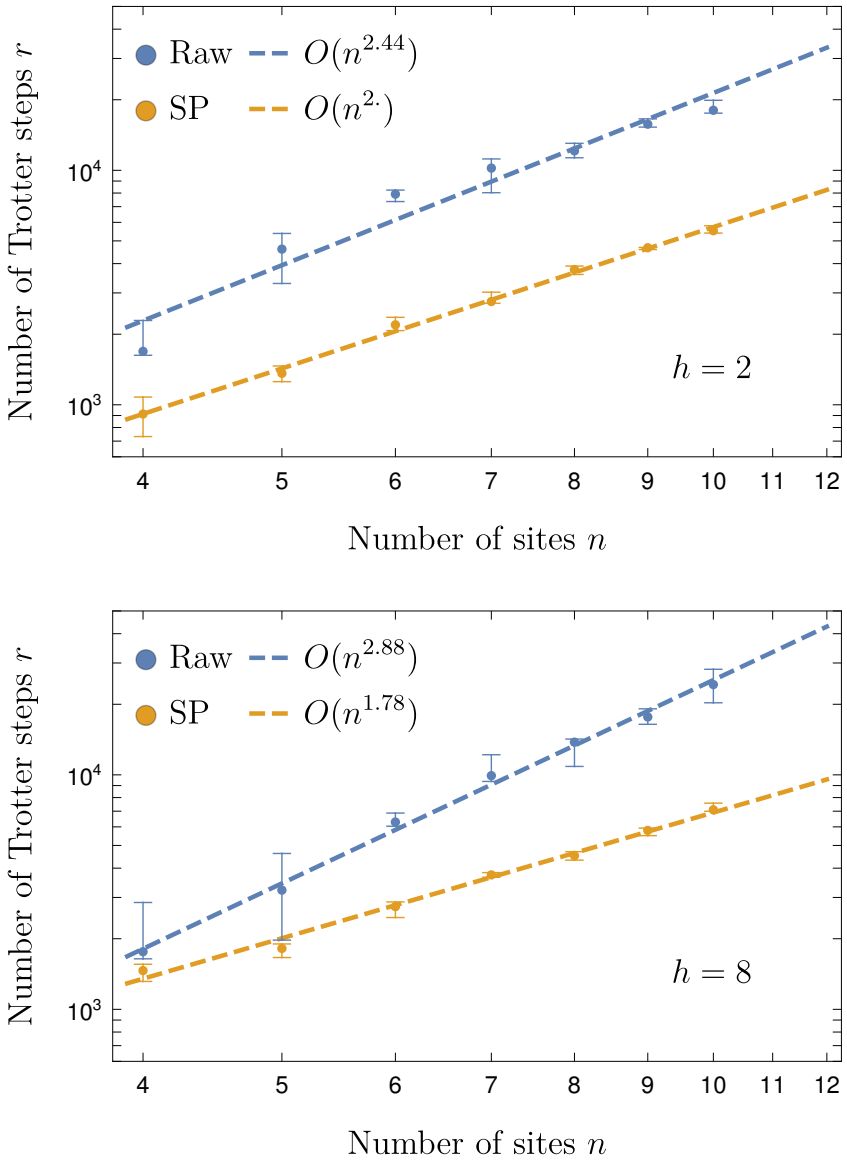

Figure 8.4: The number of Trotter steps required for the simulation of n qubits evolved under Eq. (8.40) for time $t = n$ to meet a fixed error tolerance $\varepsilon = 0.01$. We compare this Trotter number of a simulation without any symmetry protection ("Raw", blue) and a simulation with random symmetry protection ("SP", orange) at $h = 2$ (left panel) and $h = 8$ (right panel), which correspond to the system being in the ETH and the MBL respectively. The dashed lines are the linear fits of the data in the log-log scale. The simulation is repeated 100 times with different instance of the disorder h_i. The dots represent the median of the Trotter number at each n and the error bars correspond to the 25%-75% percentile region. The numerics show that symmetry protecting the simulation reduces the number of Trotter steps, and hence the gate count, by about 2 to 4 times in both the ETH and the MBL phases.

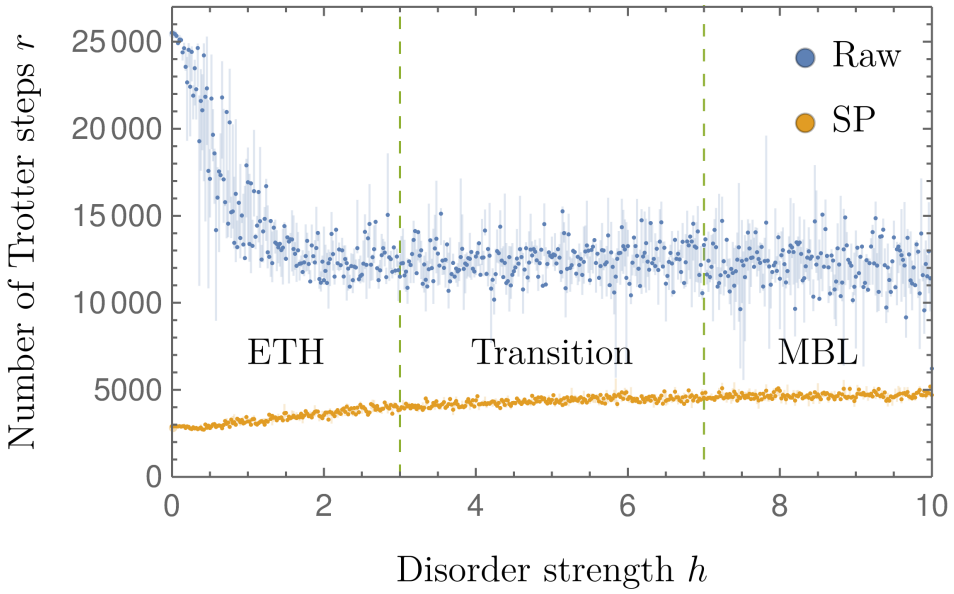

Figure 8.5: The required number of Trotter steps in simulating the Hamiltonian Eq. (8.40) of $n = 8$ qubits for time $t = n$ as a function of the disorder strength in an unprotected simulation ("Raw", blue) and in a symmetry-protected simulation ("SP", orange). Each dot represents the median Trotter number over 100 different instances of the random fields. The bars correspond to the 25%-75% percentile region.

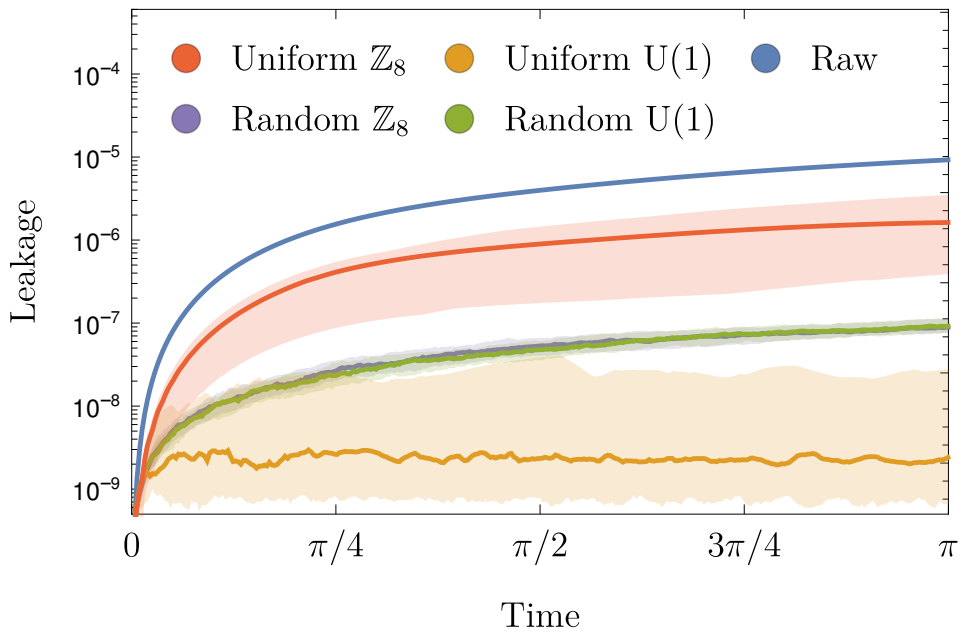

Figure 8.6: The probability for the final state to leak outside the physical subspace due to Trotter errors in simulating the Schwinger model. We consider simulations without symmetry protection (blue) and with symmetry protection under different schemes: uniform sets of transformations drawn from $\mathbb{Z}_8$ (red) and U(1) (orange) and random sets of transformations drawn from $\mathbb{Z}_8$ (purple) and U(1) (green). The purple and green areas overlap each other almost completely. The dots correspond to the median and the shaded areas correspond to the 25%-75% percentile of 100 repetitions.

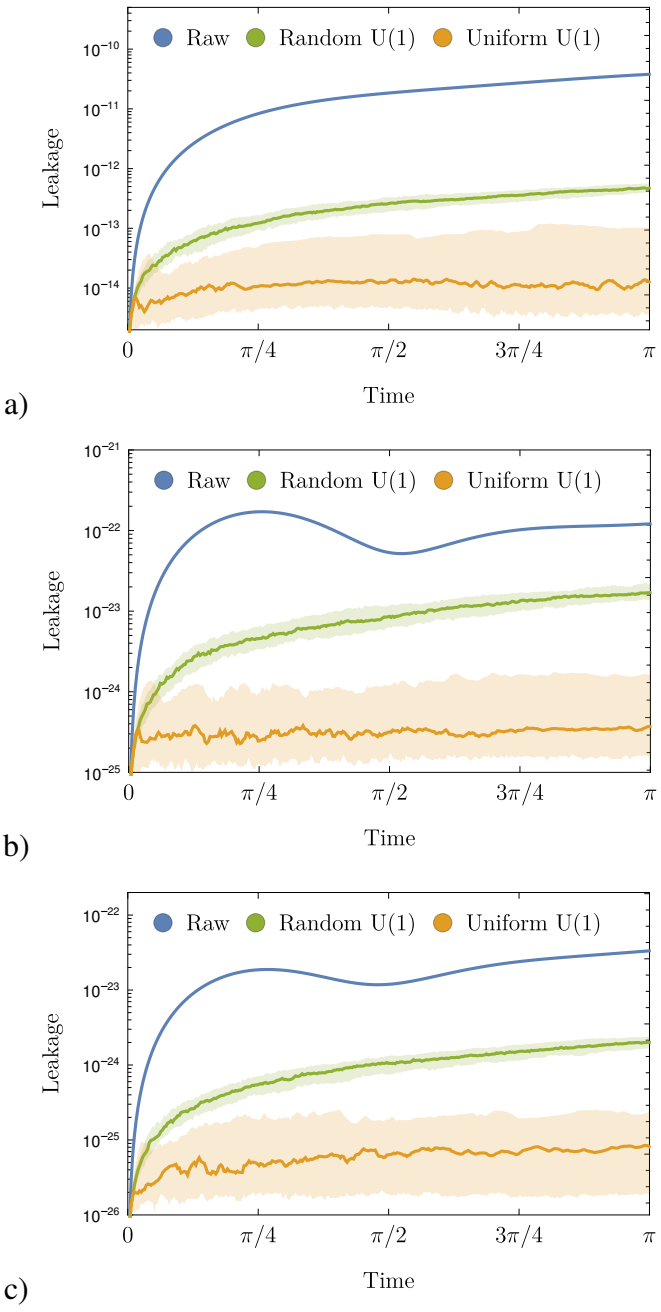

Figure 8.7: The leakage to the unphysical subspace as a function of time in simulating the Schwinger model using advanced algorithms. We consider a raw simulation (blue), a simulation protected by a random set of transformations drawn from the U(1) symmetry group (green), and a simulation protected by a uniform set of transformations (orange). The solid dots correspond to the median of 100 repetitions and the shaded area corresponds to the 25%-75% percentile.

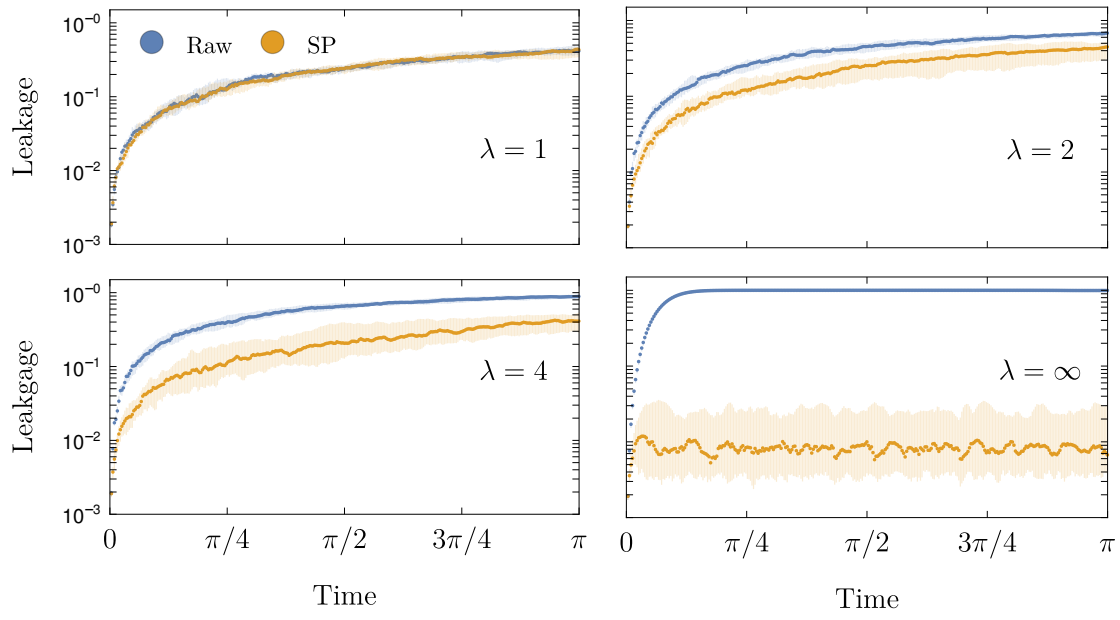

Figure 8.8: The leakage probability due to experimental noise as a function of time at different values of the correlation length λ. The simulation is repeated 100 times with different instances of the experimental noise. The solid dots represent the median of the leakage and the bars correspond to the 25%-75% percentile regions.

Chapter 9: Outlook

In this dissertation, we have demonstrated a natural relationship between digital quantum simulation, locality, and symmetry of quantum systems. In particular, the connection between quantum simulation algorithms and the Lieb-Robinson bounds that we have established in Chapter 5 suggests that one could potentially study properties of a system by merely analyzing a quantum circuit that simulates the dynamics of the system. This strategy opens another avenue for the condensed matter and atomic/molecular/optical physics communities to potentially benefit from future advances in quantum algorithms.

Conversely, the results presented in this dissertation also suggest that we can exploit advance knowledge about certain properties of a system (such as locality in Chapter 5 and symmetry in Chapter 8) to improve the performance of quantum simulation algorithms. More generally, it is conceivable that the optimal algorithm would always make the best use of information available prior to the simulation. However, it is not always clear how certain information, such as the results from shorter-time simulations or the statistics of past measurements, can be systematically incorporated to speed up digital quantum simulation.

Besides this general connection between quantum simulation algorithms and properties of quantum systems, there are several interesting open questions specific to locality

and symmetry of quantum systems. For example, since the Lieb-Robinson bound derived from the HHKL algorithm in Chapter 5 is not tight, does that imply the HHKL algorithm is not optimal in simulating power-law decaying interactions? If so, can we make use of a tighter Lieb-Robinson bound to invent a more efficient quantum simulation algorithm?

Concerning the fast state-transfer protocol described in Chapter 3, we do not know whether there exists a time-*independent* power-law Hamiltonian that propagates information at the same speed as our protocol does. Such a Hamiltonian would enable observation of fast information propagation in existing experimental platforms where arbitrary time-dependent control is often not available. On the other hand, the lack of such a Hamiltonian would imply a more stringent speed limit for time-independent Hamiltonians than the one given by the Lieb-Robinson bound. Such a speed limit may in turn imply that the effective time-independent Floquet Hamiltonians constructed in Refs. [74, 75, 137] for periodically driven, power-law interacting systems would correctly generate the dynamics of local observables even in the regime $d < \alpha < 2d$, closing the gap in our understanding mentioned in Chapter 3. These observations emphasize the need for studies, in a similar spirit to Chapter 4, of fundamental speed limits for time-independent Hamiltonians and, more generally, systems under various physical constraints.

Another interesting open question is whether our optimal protocol can be generalized to the regime $0 \leq \alpha \leq d$, where there are still substantial gaps between the Lieb-Robinson bounds and achievable protocols [24, 30, 193, 194, 195]. The bounds suggest that, in addition to the distance, the information-propagation time also depends on the total number of sites on the lattice. Consequently, we would expect an optimal protocol to make use of all sites on the lattice, including those that are far from both the source

and the target of the propagation. We consider such a generalization an important future direction.

Concerning the symmetry-protected simulations described in Chapter 8, an immediate future direction is to generalize the analysis in this chapter to more advanced quantum simulation algorithms, such as the higher-order Suzuki-Trotter formulas [160], the truncated Taylor series [8], or qubitization [11]. We emphasize that our approach induces destructive interference between the errors from different steps of the simulation and, therefore, should suppress errors that violate the symmetries of the target system, regardless of the sources of the errors. However, the optimal choice of the symmetry transformations depends on the exact error structure in each step of the simulation. Since the error structures of more advanced algorithms are typically more complicated than the first-order Trotterization, it is more difficult to infer the set of symmetry transformations that optimally protects the simulation. Nevertheless, extensive analytical and numerical studies of the effectiveness of the technique for protecting these advanced algorithms, especially when applied to the simulations of various physically relevant systems, such as the lattice field theories [180, 181, 182], or the electronic structures [158, 196, 197, 198], would be useful for the long-term development of digital quantum simulation.

When the error structure of the algorithm is not readily available, an alternative promising approach for optimizing the set of symmetry transformations is to parameterize the transformations, variationally minimize the error of the first few simulation steps, and apply the same set of transformations repeatedly for the rest of the simulation [199]. Understanding when such a variational approach can suppress the error in a long simulation could provide a path towards a scalable symmetry protection with a minimal

calculation overhead.

In addition, our analysis in Chapter 8 focuses primarily on the error of the simulation algorithm under the symmetry protection in the full Hilbert space. It would be interesting to, for example, build upon the recent result of Ref. [200] and analyze the symmetry-protected simulation error in a low-energy subspace. The main ingredient of Ref. [200] is an analysis of how evolutions under the terms of the Hamiltonian take the system from one eigen-subspace to another. We can potentially generalize their approach to analyze the leakage between symmetry sectors and subsequently study the effects of the symmetry protection on such leakage.

Appendices

Appendix A: Supplemental Material for Chapter 2

In this appendix, we provide a rigorous proof of Theorem 1 in the main text (Appendix A.1) and details on the applications of the bound to connected correlators, topologically ordered states, and simulations of local observables (Appendix A.2).

A.1 Proof of Theorem 1

In this section, we provide a rigorous proof of Theorem 11. We first summarize the lemmas we use in the proof of the theorem, followed by the proofs of the lemmas in Appendices A.1.1 to A.1.3.

For convenience, we first recall the definitions from the main text. We consider a d-dimensional lattice of qubits Λ and, acting on this lattice, a two-body power-law Hamiltonian $H(t)$ with exponent α. Specifically, we assume $H(t) = \sum_{i,j\in\Lambda} h_{ij}(t)$ is a sum of two-body terms h_{ij} supported on sites i, j such that $\|h_{ij}(t)\| \leq 1/\text{dist}(i,j)^\alpha$ for all $i \neq j$, where $\|\cdot\|$ is the operator norm and $\text{dist}(i,j)$ is the distance between i, j. In this chapter, we assume $2d < \alpha < 2d + 1$.

We use $\mathcal{L}$ to denote the Liouvillian corresponding to the Hamiltonian H, i.e. $\mathcal{L}\,|O) \equiv i\,|[H, O])$ for all operators O, and use $e^{\mathcal{L}t}\,|O) \equiv |O(t))$ to denote the time evolved version of the operator O. Similarly to the main text, we use $\mathbb{P}_r^{(i)}\,|O)$ to denote the projection of

O onto sites that are at least a distance r from site i. In particular, if i is the origin of the lattice, we may also drop the superscript i and simply write $\mathbb{P}_r$ for brevity.

Given a unit-norm operator O initially supported at the origin, $\mathbb{P}_r e^{\mathcal{L}t}|O)$ provides the fraction of the time-evolved version of the operator O that is supported at least a distance r from the origin at time t. The identity [22]

$$\frac{1}{2} \leq \frac{\|\mathbb{P}_r e^{\mathcal{L}t}|O)\|}{\sup_A \|[A, e^{\mathcal{L}t}O]\|} \leq 2, \tag{A.1}$$

where the supremum is taken over all unit-norm operators A supported at least a distance r from O, establishes the equivalence between the projector and the unequal-time commutator commonly used in the Lieb-Robinson literature.

Theorem 11. *For any $\alpha \in (2d, 2d+1)$ and $\varepsilon \in \left(0, \frac{(\alpha-2d)^2}{(\alpha-2d)^2+\alpha-d}\right)$, there exist constants $c, C_1, C_2 \geq 0$ such that*

$$\|\mathbb{P}_r e^{\mathcal{L}t}|O)\| \leq C_1 \left(\frac{t}{r^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\frac{\varepsilon}{2}} + C_2 \frac{t}{r^{\alpha-d}} \tag{A.2}$$

holds for all $t \leq c r^{\alpha-2d-\varepsilon}$.

Our strategy is to divide the terms of the Hamiltonian by their interaction range and prove a Lieb-Robinson-like bound recursively for each range. Specifically, let $\ell_0 = 0$ and $\ell_k \equiv L^k$ for $k = 1, 2, \ldots, n$, where $L > 1$ to be chosen later,

$$n = \left\lfloor \frac{1}{\log L} \log \left[r \left(\frac{t}{r^{\alpha-2d}} \right)^{\eta} \right] \right\rfloor, \tag{A.3}$$

and $\eta \in (0, \frac{1}{\alpha - d})$ is an arbitrary small constant. For our convenience, we set $\ell_{n+1} = r_*$, where r_* is the diameter of the lattice. We then divide the Hamiltonian into $H = \sum_{k=1}^{n+1} V_k$, where $V_k = \sum_{i,j:\ell_{k-1} < \text{dist}(i,j) \le \ell_k} h_{ij}$ consists of terms h_{ij} such that the distance between i, j is between ℓ_{k-1} and ℓ_k. We also use $H_k = \sum_{j=1}^{k} V_k$ to denote the sum of interactions whose lengths are at most ℓ_k and $\mathcal{L}_k = i[H_k, \cdot]$ are the corresponding Liouvillians. Note that $H_{n+1} = H$ contains every interaction of the Hamiltonian.

We start with a standard Lieb-Robinson bound for H_1 [18, 34], i.e.

$$\|\mathbb{P}_r e^{\mathcal{L}_1 t}(O)\| \le \exp\left[\frac{v_1 t - r}{\ell_1}\right], \tag{A.4}$$

where $v_1 = 4e\tau\ell_1$ is proportional to ℓ_1 and $\tau = \max_i \sum_{j \in \Lambda, j \ne i} 1/\text{dist}(i,j)^\alpha$ is a constant for all $\alpha > d$, and recursively prove bounds for $H_2, H_3, \ldots, H_n$ using the following lemma:

Lemma 12. *Suppose for $\ell_k \ge 1$, we have*

$$\|\mathbb{P}_r e^{\mathcal{L}_k t}(O)\| \le \exp\left[\frac{v_k t - r}{\ell_k}\right], \tag{A.5}$$

for some unit-norm operator O supported at the origin. Then for $\ell_{k+1} > \ell_k$, we have

$$\|\mathbb{P}_r e^{\mathcal{L}_{k+1} t}(O)\| \le \exp\left[\frac{v_{k+1} t - r}{\ell_{k+1}}\right]. \tag{A.6}$$

where

$$v_{k+1} = \xi \log(r_*) v_k + \nu\lambda \frac{\ell_{k+1}^{2d+1}}{\ell_k^{\alpha}} \tag{A.7}$$

and ξ, ν, λ are constants that may depend only on d.

Note that each of the bounds in the series has a logarithmic dependence on the diameter r_* of the lattice. We later show that this dependence on r_* can be replaced by a similar logarithmic dependence on r, leading to a logarithmic correction in the light cone. After applying Lemma 12 $n-1$ times, we arrive at a bound for the evolution under H_n:

$$\|\mathbb{P}_r e^{\mathcal{L}_n t} |O)\| \leq \exp\left[\frac{v_n t - r}{\ell_n}\right], \tag{A.8}$$

where

$$v_n = x^{n-1}(v_1 - L^{2d+1}\nu\lambda) + x^{n-1}L^{2d+1}\nu\lambda\left[1 + \frac{L^{2d+1-\alpha}}{x} + \ldots \left(\frac{L^{2d+1-\alpha}}{x}\right)^{n-1}\right] \tag{A.9}$$

and $x \equiv \xi \log r_*$. We now choose $L = x^{1/(2d+1-\alpha)}$ so that

$$v_n = x^{n-1}[v_1 + (n-1)L^{2d+1}\nu\lambda]. \tag{A.10}$$

At this point, we have a bound for the evolution under H_n, which contains most terms of the Hamiltonian except for those with range larger than ℓ_n. With the value of n in Eq. (A.3), we eventually show that the bound Eq. (A.8) has the desired light cone

$t \gtrsim r/v_n \sim r^{\alpha-2d}$.

Next, we add the remaining long-range interactions in $H - H_n$, i.e. those with range larger than ℓ_n, to the bound. The result is the following lemma, which we prove in Appendix A.1.2.

Lemma 13. *Given any $\varepsilon > 0$, there exist constants C, c, κ, δ such that*

$$\|\mathbb{P}_r e^{\mathcal{L}t}|O)\| \leq C \log^{\kappa} r_* \left(\frac{t}{r^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\varepsilon} \tag{A.11}$$

holds for all $t \leq c r^{\alpha-2d-\varepsilon}/\log^{\delta} r_*$.

The bound at this point still has an undesirable feature: it depends on the size of the lattice r_*. Finally, we show in Appendix A.1.3 that we can remove this dependence on r_* at the cost of adding additional terms to the bound. The result is Theorem 11 presented in the main text.

A.1.1 Proof of Lemma 12

In this section, we prove Lemma 12.

Proof. For simplicity, let $V \equiv V_{k+1} = H_{k+1} - H_k$ in this section. We shall move into the interaction picture of H_k and write the time evolution under H_{k+1} as a product

$$\mathcal{T}\exp\left(-i\int_0^t ds\, H_{k+1}(s)\right) = \mathcal{T}\exp\left(-i\int_0^t ds\, e^{\mathcal{L}_k s}V\right)\cdot\mathcal{T}\exp\left(-i\int_0^t ds\, H_k(s)\right) \tag{A.12}$$

of an evolution under H_k, for which Eq. (A.5) applies, and an evolution $e^{\mathcal{L}_I t}$ under the $V_I(t) = e^{\mathcal{L}_k t} V$.

We decompose every term h_{ij} in V into a sum of products of two single-site operators $u_i^{(\mu)}$:

$$h_{ij} = \sum_{\mu} J_{ij}^{(\mu)} u_i^{(\mu)} u_j^{(\mu)}, \tag{A.13}$$

where $u_i^{(\mu)}$ have unit norms, $J_{ij}^{(\mu)}$ are nonnegative, and $\sum_\mu J_{ij}^{(\mu)} \leq 1/\mathrm{dist}(i,j)^\alpha$. In doing so, we can reduce the evolution of h_{ij} into the evolutions of single-site operators $u_i^{(\mu)}$:

$$e^{\mathcal{L}_k t} h_{ij} = \sum_{\mu} J_{ij}^{(\mu)} \left[e^{\mathcal{L}_k t} u_i^{(\mu)} \right] \left[e^{\mathcal{L}_k t} u_j^{(\mu)} \right]. \tag{A.14}$$

We then pick a parameter $R \geq \ell_k$ and divide the lattice around i into shells of width R. Specifically, let $\mathcal{B}_r^{(i)}$ denote the ball of radius r centered on i. Let $\mathcal{S}_r^{(i)} = \mathcal{B}_r^{(i)} \setminus \mathcal{B}_{r-R}^{(i)}$ denote the shell of inner radius $r - R$ and outer radius r centered on i. For each μ, we have

$$e^{\mathcal{L}_k t} u_i^{(\mu)} = \left[(\mathbb{I} - \mathbb{P}_R^{(i)}) + (\mathbb{P}_R^{(i)} - \mathbb{P}_{2R}^{(i)}) + (\mathbb{P}_{2R}^{(i)} - \mathbb{P}_{3R}^{(i)}) + \dots \right] e^{\mathcal{L}_k t} u_i^{(\mu)} \equiv \sum_{q=0}^{\infty} u_{i,q}^{(\mu)}(t), \tag{A.15}$$

where the distance in the subscript of the projectors is with respect to i and $u_{i,q}^{(\mu)}$ is supported on $\mathcal{B}_{(q+1)R}^{(i)}$ for $q = 0, 1, 2, \dots$.

Using Eq. (A.5) and the triangle inequality, we can show that

$$\|u_{i,q}^{(\mu)}(t)\| \leq \|\mathbb{P}_{qR}^{(i)} e^{\mathcal{L}_k t} u_i^{(\mu)}\| + \|\mathbb{P}_{(q+1)R}^{(i)} e^{\mathcal{L}_k t} u_i^{(\mu)}\| \tag{A.16}$$

$$\leq \exp\left(\frac{v_k t - qR}{\ell_k}\right) + \exp\left(\frac{v_k t - (q+1)R}{\ell_k}\right). \tag{A.17}$$

Choosing $R \geq v_k t$ and $R \geq (1+\varepsilon)\ell_k$ for some positive constant ε, we have

$$\|u_{i,q}^{(\mu)}(t)\| \leq e^{\frac{-(q-1)R}{\ell_k}} + e^{\frac{-qR}{\ell_k}} \leq e^{-(q-1)(1+\varepsilon)} + e^{-q(1+\varepsilon)} \leq (1+e^{1+\varepsilon}) e^{-q(1+\varepsilon)} \tag{A.18}$$

for all $q = 0, 1, 2, \ldots$. By combining the two legs of h_{ij} together, we arrive at a decomposition $e^{\mathcal{L}_k t} h_{ij} = \sum_{p,q} w_{i,p;j,q}(t)$, where $w_{i,p;j,q}(t) = \sum_\mu J_{ij}^{(\mu)} u_{i,p}^{(\mu)}(t) u_{j,q}^{(\mu)}(t)$ and

$$\|w_{i,p;j,q}(t)\| \leq \frac{(1+e^{1+\varepsilon})^2}{\text{dist}(i,j)^\alpha} e^{-(p+q)(1+\varepsilon)}. \tag{A.19}$$

Next, we divide the lattice into complementary hypercubes of length R. We shall prove that $V_I(t)$ actually consists of exponentially decaying interactions between hypercubes. We shall index the hypercubes by their centers, i.e. $\mathcal{C}_x$ denotes the hypercube center at x. Given x, y as the centers of two hypercubes,

$$\tilde{h}_{xy}(t) \equiv \sum_{\substack{i,j,p,q \\ \mathcal{B}_{(p+1)R}^{(i)} \cap \mathcal{C}_x \neq \varnothing \\ \mathcal{B}_{(q+1)R}^{(j)} \cap \mathcal{C}_y \neq \varnothing}} w_{i,p;j,q}(t) \tag{A.20}$$

defines the effective interaction between the cubes $\mathcal{C}_x$ and $\mathcal{C}_y$. Note that $\sum_{x,y} \tilde{h}_{xy} \neq V_I$ because some $w_{i,p;j,q}$ might be double counted. The conditions $\mathcal{B}_{(p+1)R}^{(i)} \cap \mathcal{C}_x \neq \varnothing$ and

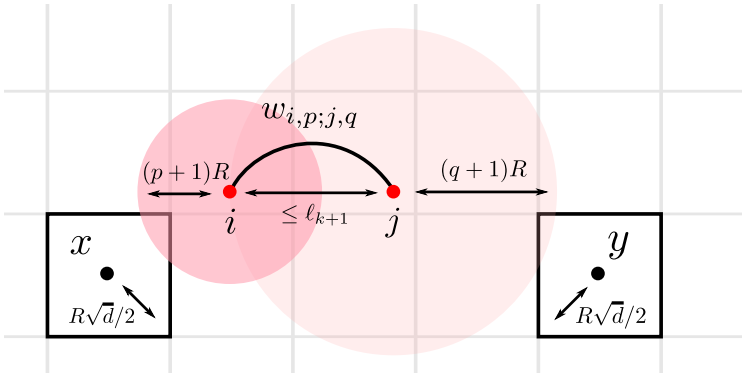

Figure A.1: The effective interaction between two hypercubes $\mathcal{C}_x$ and $\mathcal{C}_y$ comes from the terms $w_{i,p;j,q}$ whose support (the shaded area) overlaps with the cubes.

$\mathcal{B}^{(j)}_{(q+1)R} \cap C_y \neq \varnothing$ ensure that we account for all terms $w_{i,p;j,q}(t)$ whose support might overlap with the cubes $\mathcal{C}_x, \mathcal{C}_y$ (Fig. A.1). These conditions, together with $\text{dist}(i,j) \leq \ell_{k+1}$, can be relaxed to

1. $\text{dist}(i,x) \leq (p+1)R + R\frac{\sqrt{d}}{2}$,
2. $\text{dist}(j,y) \leq (q+1)R + R\frac{\sqrt{d}}{2}$, and
3. $\text{dist}(x,y) \leq (p+1)R + R\frac{\sqrt{d}}{2} + \ell_{k+1} + (q+1)R + R\frac{\sqrt{d}}{2}$,

where $(p+1)R$ and $(q+1)R$ are the radii of the balls around i and j, $R\sqrt{d}/2$ is the maximum distance between the center and the corner of a hypercube, and the middle term ℓ_{k+1} comes from the maximum distance between i and j.

We bound the norm of $\tilde{h}_{xy}(t)$ using the triangle inequality and relax the conditions for i, j, p, q as we discussed above:

$$\|\tilde{h}_{xy}(t)\| \leq \sum_{\substack{p,q,i,j \\ (1),(2),(3)}} \|w_{i,p;j,q}(t)\| \leq \sum_{\substack{p,q,i,j \\ (1),(2),(3)}} \frac{(1+e^{1+\varepsilon})^2}{\text{dist}(i,j)^\alpha} e^{-(p+q)(1+\varepsilon)}, \tag{A.21}$$

where the subscript $(1), (2), (3)$ of the sum refers to the three conditions above, respec-

tively. Since $\text{dist}(i,j) \geq \ell_k$, we can simplify the bound and carry out the sums over i, j:

$$\|\tilde{h}_{ij}(t)\| \leq \frac{(1+e^{1+\varepsilon})^2}{\ell_k^\alpha} \sum_{\substack{p,q,i,j \\ (1),(2),(3)}} e^{-(p+q)(1+\varepsilon)} \tag{A.22}$$

$$\leq \frac{(1+e^{1+\varepsilon})^2}{\ell_k^\alpha} \sum_{\substack{p,q \\ (3)}} 4^d \left(R + pR + R\frac{\sqrt{d}}{2}\right)^d \left(R + qR + R\frac{\sqrt{d}}{2}\right)^d e^{-(p+q)(1+\varepsilon)} \tag{A.23}$$

$$= \frac{(1+e^{1+\varepsilon})^2}{\ell_k^\alpha} (2R)^{2d} \sum_{\substack{p,q \\ (3)}} \left(p + 1 + \frac{\sqrt{d}}{2}\right)^d \left(q + 1 + \frac{\sqrt{d}}{2}\right)^d e^{-(p+q)(1+\varepsilon)}. \tag{A.24}$$

We then use the following identity to simplify the expression: For every $\varepsilon > 0$,

$$x^d \leq g_\varepsilon e^{\varepsilon x} \tag{A.25}$$

holds for all $x \geq 0$, where $g_\varepsilon = d!/\varepsilon^d$. Therefore, we can bound

$$\left(p + 1 + \frac{\sqrt{d}}{2}\right)^d \leq g_\varepsilon e^{\varepsilon + \varepsilon\frac{\sqrt{d}}{2}} e^{\varepsilon p}. \tag{A.26}$$

Substituting back to the earlier equation, we have

$$\|\tilde{h}_{xy}(t)\| \leq g_\varepsilon^2 e^{2\varepsilon + 2\varepsilon\frac{\sqrt{d}}{2}} \frac{(1+e^{1+\varepsilon})^2}{\ell_k^\alpha} (2R)^{2d} \sum_{\substack{p,q \\ (3)}} e^{-(p+q)(1+\varepsilon-\varepsilon)} \leq \frac{\tilde{g}_\varepsilon}{\ell_k^\alpha} R^{2d} \sum_{\substack{p,q \\ (3)}} e^{-(p+q)}, \tag{A.27}$$

where $\tilde{g}_\varepsilon$ absorbs all constants that depend only on ε and d. Recall that condition (3) is

equivalent to

$$p + q \geq \frac{\text{dist}(x, y)}{R} - 2 - \sqrt{d} - \frac{\ell_{k+1}}{R} \equiv a. \tag{A.28}$$

We consider two cases. For $q \geq a$, the sum over p can be taken from 0 to ∞:

$$\begin{aligned}\frac{\tilde{g}_\varepsilon}{\ell_k^\alpha} R^{2d} \sum_{q \geq a} \sum_{p \geq 0} e^{-(p+q)} &\leq \frac{\tilde{g}_\varepsilon}{\ell_k^\alpha} R^{2d} e^{-a+1} \sum_{q \geq 0} \sum_{p \geq 0} e^{-(p+q)} = \frac{e^3 \tilde{g}_\varepsilon}{(e-1)^2 \ell_k^\alpha} R^{2d} e^{-a} \\ &= \frac{e^3 \tilde{g}_\varepsilon}{(e-1)^2} e^{2+\sqrt{d}} e^{\frac{\ell_{k+1}}{R}} \frac{R^{2d}}{\ell_k^\alpha} e^{-\frac{\text{dist}(i,j)}{R}}.\end{aligned} \tag{A.29}$$

For $q < a$, we sum over $p \geq a - q$:

$$\frac{\tilde{g}_\varepsilon}{\ell_k^\alpha} R^{2d} \sum_{q < a} \sum_{p \geq a-q} e^{-(p+q)} \leq \frac{e^2}{e-1} \frac{\tilde{g}_\varepsilon}{\ell_k^\alpha} R^{2d} g_\varepsilon e^{-(1-\varepsilon)a} \leq \frac{e^2 \tilde{g}_\varepsilon g_\varepsilon e^{2+\sqrt{d}}}{e-1} e^{\frac{\ell_{k+1}}{R}} \frac{R^{2d}}{\ell_k^\alpha} e^{-(1-\varepsilon)\frac{\text{dist}(i,j)}{R}}, \tag{A.30}$$

where we have used the identity Eq. (A.25) again with $d \geq 1$ and $\varepsilon > 0$ having the same value as before.

Combining Eqs. (A.29) and (A.30), we have

$$\|\tilde{h}_{xy}(t)\| \leq \underbrace{\tilde{g}_\varepsilon \frac{e^2}{e-1} e^{2+\sqrt{d}} \left(\frac{e}{e-1} + g_\varepsilon \right) e^{\frac{\ell_{k+1}}{R}} \frac{R^{2d}}{\ell_k^\alpha}}_{=\mathcal{E}_0} e^{-(1-\varepsilon)\frac{\text{dist}(x,y)}{R}}. \tag{A.31}$$

Note that $\frac{\text{dist}(x,y)}{R}$ is the rescaled distance between the hypercubes $\mathcal{C}_x$ and $\mathcal{C}_y$. Therefore, the interaction between the hypercubes decays exponentially with the rescaled distance between them. Using the standard Lieb-Robinson bound for exponentially decaying in-

teractions, there exists a constant ν such that

$$\|\mathbb{P}_r e^{\mathcal{L}_I t}|O)\| \leq \exp\left(\nu \mathcal{E}_0 t - \frac{(1-\varepsilon)r}{R}\right) \tag{A.32}$$

for any unit-norm operator O supported on a single hypercube (including operators supported on single sites.) We now choose $R = (1-\varepsilon)\ell_{k+1}$ and rewrite

$$\mathcal{E}_0 = \underbrace{\tilde{g}_\varepsilon \frac{e^2}{e-1} e^{2+\sqrt{d}} \left(\frac{e}{e-1} + g_\varepsilon\right) e^{\frac{1}{1-\varepsilon}} (1-\varepsilon)^{2d}}_{\equiv \lambda} \frac{\ell_{k+1}^{2d}}{\ell_k^\alpha}, \tag{A.33}$$

where the constant λ depends only on ε and d. Plugging this expression into the earlier bound, we get

$$\|\mathbb{P}_r e^{\mathcal{L}_I t}|O)\| \leq \exp\left(\frac{\nu\lambda \frac{\ell_{k+1}^{2d+1}}{\ell_k^\alpha} t - r}{\ell_{k+1}}\right) = \exp\left(\frac{\Delta v\, t - r}{\ell_{k+1}}\right), \tag{A.34}$$

where

$$\Delta v \equiv \nu\lambda \frac{\ell_{k+1}^{2d+1}}{\ell_k^\alpha}. \tag{A.35}$$

Note that we assume $R = (1-\varepsilon)\ell_{k+1} \geq (1+\varepsilon)\ell_k$. A constant ε satisfying this condition exists as long as $\ell_{k+1} > \ell_k$.

Next, we use the following lemma to "merge" this bound for $e^{\mathcal{L}_I t}$ with the bound in Eq. (A.5) for $e^{\mathcal{L}_k t}$.

Lemma 14. *Let H_1, H_2 be two possibly time-dependent Hamiltonians and $\mathcal{L}_1, \mathcal{L}_2$ be the*

corresponding Liouvillians. Suppose that for all unit-norm, single-site operators O and for all times $t \leq \Delta t$ for some Δt,

$$\|\mathbb{P}_r e^{\mathcal{L}_1 t} |O)\| \leq c_1 r^{\xi_1} e^{\frac{v_1 t - r}{\ell_1}}, \tag{A.36}$$

$$\|\mathbb{P}_r e^{\mathcal{L}_2 t} |O)\| \leq c_2 r^{\xi_2} e^{\frac{v_2 t - r}{\ell_2}}, \tag{A.37}$$

for some $\ell_2 \geq \ell_1$ and $c_1, c_2 \geq 1; \xi_1, \xi_2 \geq 0$ are constants. We have

$$\|\mathbb{P}_r e^{\mathcal{L}_2 t} e^{\mathcal{L}_1 t} |O)\| \leq 2^{d+5} c_1 c_2 r^{\xi_1+\xi_2+d+1} e^{\frac{(v_1+v_2)t - r}{\ell_2}}, \tag{A.38}$$

for all $t \leq \Delta t$.

We prove Lemma 14 in Appendix A.1.1.1. Using the lemma, we obtain a bound for the evolution under H_{k+1}:

$$\|\mathbb{P}_r e^{\mathcal{L}_{k+1} t} |O)\| = \|\mathbb{P}_r e^{\mathcal{L}_I t} e^{\mathcal{L}_k t} |O)\| \leq 2^{d+5} r^{d+1} e^{\frac{(v_k + \Delta v)t - r}{\ell_{k+1}}}. \tag{A.39}$$

However, because we assume $v_k t \leq R$ in deriving Eq. (A.34), Eq. (A.39) is only valid for small time $t \leq (1-\varepsilon)\ell_{k+1}/v_k \equiv \Delta t$. To extend the bound to all time, we use a corollary of Lemma 14:

Corollary 4. *Suppose we have a single-site, unit-norm operator O, a Hamiltonian H with a corresponding Liouvillian $\mathcal{L}$, a constant Δt, and*

$$\|\mathbb{P}_r e^{\mathcal{L} t} |O)\| \leq c_0 r^{\xi_0} e^{(vt - r)/\ell} \tag{A.40}$$

holds for all $t \le \Delta t$. Then, for all $t \le 2^k \Delta t$ for any $k \in \mathbb{N}$, we have

$$\|\mathbb{P}_r e^{\mathcal{L}t} |O)\| \le c_k r^{\xi_k} e^{(vt-r)/\ell}, \tag{A.41}$$

where $c_k = 2^{(d+5)(2^k-1)} c_0^{2^k}$, $\xi_k = (2^k-1)(d+1) + 2^k \xi_0$ are constants. In particular,

$$\|\mathbb{P}_r e^{\mathcal{L}t} |O)\| \le e^{\frac{\chi t}{\Delta t} + \frac{vt-r}{\ell}}, \tag{A.42}$$

where $\chi = 2[\log(2^{d+5} c_0) + (d+1+\xi_0)\log r]$, holds for all time t.

We prove the corollary in Appendix A.1.1.2. Using the corollary, we can extend Eq. (A.39) to a bound for all time:

$$\begin{aligned} \|\mathbb{P}_r e^{\mathcal{L}_{k+1}t} |O)\| &\le \exp\left(\frac{\chi t}{\Delta t} + \frac{(v_k + \Delta v)t - r}{\ell_{k+1}}\right) && \text{(A.43)} \\ &\le \exp\left(\frac{\chi_* v_k t}{(1-\varepsilon)\ell_{k+1}} + \frac{(v_k + \Delta v)t - r}{\ell_{k+1}}\right) && \text{(A.44)} \\ &= \exp\left(\frac{v_{k+1} t - r}{\ell_{k+1}}\right), && \text{(A.45)} \end{aligned}$$

where we have upper bounded χ by $\chi_* = 4(d+5)\log 2 + 4(d+1)\log r_*$, $r_* \ge r$ is the diameter of the lattice, and

$$v_{k+1} = \left(\frac{\chi_*}{1-\varepsilon} + 1\right) v_k + \nu\lambda \frac{\ell_{k+1}^{2d+1}}{\ell_k^\alpha} \le 4(4d+13)\log(r_*) v_k + \nu\lambda \frac{\ell_{k+1}^{2d+1}}{\ell_k^\alpha}. \tag{A.46}$$

Here, we have assumed that $r_* \ge 2$ and $\varepsilon \le 1/2$ so that $1/(1-\varepsilon) \le 2$, $\chi_* \le 4(2d+6)\log r_*$, and $1 \le 4 \log r_*$. Therefore, Lemma 12 holds with $\xi = 4(4d+13)$. □

A.1.1.1 Proof of Lemma 14

In this section, we prove Lemma 14.

Proof. The bound is trivial for $r < vt$, where $v = v_1 + v_2$. Therefore, we will consider $r \geq vt$ in the rest of the proof.

The strategy is to apply Eqs. (A.36) and (A.37) consecutively. A technical difficulty comes from the fact that after the first evolution $e^{\mathcal{L}_1 t}$, the operator has spread to more than one site. Therefore, we cannot directly apply Eq. (A.37), which assumes that the operator is single-site. Instead, we need to use [25, Lemma 4] to extend the bound for sing-site operators to multi-site operators. In particular, given the assumed bound Eq. (A.37) and an unit-norm operator O_X supported on a ball X of radius $x \leq r$, we have

$$\|\mathbb{P}_r e^{\mathcal{L}_2 t}(O_X)\| \leq \frac{9}{2}|X| c_2 r^{\xi_2} e^{(v_2 t - r + x)/\ell_2}. \tag{A.47}$$

With that in mind, we divide the lattice into:

1. A ball of radius $v_1 t$ around the origin,
2. Shells of inner radius $v_1 t + (q-1)\ell_1$ and outer radius $v_1 t + q\ell_1$ for $q = 1, \ldots, \frac{r - v_1 t}{\ell_1}$,
3. The rest of the lattice, i.e. sites at least a distance r from the origin.

We then project $e^{\mathcal{L}_1 t}|O)$ into these regions:

$$e^{\mathcal{L}_1 t}|O) = \left[(\mathbb{I} - \mathbb{P}_{v_1 t}) + \sum_{q=1}^{(r-v_1 t)/\ell_1} (\mathbb{P}_{v_1 t+(q-1)\ell_1} - \mathbb{P}_{v_1 t+q\ell_1}) + \mathbb{P}_r\right] e^{\mathcal{L}_1 t}|O) \tag{A.48}$$

$$\equiv |O_0) + \sum_{q=1}^{q_*} |O_q) + |O_*)\,, \tag{A.49}$$

where $q_* = (r - v_1 t)/\ell_1$. We then apply the other evolution, i.e. $e^{\mathcal{L}_2 t}$, on each term of the above expansion.

First, we consider $|O_0)$, which has norm at most three and is supported on at most $(2v_1 t)^d = (2v_1 t)^d \leq (2r)^d$ sites that are at least a distance $r - v_1 t$ from the outside. Using the assumed bound, we have

$$\|\mathbb{P}_r e^{\mathcal{L}_2 t}|O_0)\| \leq \frac{27}{2}(2r)^d c_2 r^{\xi_2} e^{(v_2 t - r + v_1 t)/\ell_2} = \frac{27}{2} 2^d c_2 r^{\xi_2 + d} e^{(vt-r)/\ell_2}. \tag{A.50}$$

Next, we consider $|O_*)$. Because $\|O_*\| \leq c_1 r^{\xi_1} e^{(v_1 t - r)/\ell_1}$,

$$\|\mathbb{P}_r e^{\mathcal{L}_2 t}|O_*)\| \leq 2\|O_*\| \leq 2c_1 r^{\xi_1} e^{(v_1 t - r)/\ell_1} \tag{A.51}$$

$$\leq 2c_1 c_2 r^{\xi_1+\xi_2} e^{(vt-r)/\ell_1} \leq 2c_1 c_2 r^{\xi_1+\xi_2} e^{(vt-r)/\ell_2}. \tag{A.52}$$

Finally, we consider $|O_q)$. Note that O_q is supported on a ball of volume at most $2^d(v_1 t + q\ell_1)^d \leq (2r)^d$, $\|O_q\| \leq (1+e)c_1 r^{\xi_1} e^{-q}$ and the distance between O_q and $\mathbb{P}_r$ is

$r - v_1 t - q\ell_1 \leq r$. Therefore, we have

$$\sum_q \|\mathbb{P}_r e^{\mathcal{L}_2 t} |O_q)\| \leq \sum_q \frac{9}{2}(2r)^d \|O_q\| c_2 r^{\xi_2} e^{(v_2 t - (r - v_1 t - q\ell_1))/\ell_2} \tag{A.53}$$

$$\leq \sum_q \frac{9}{2}(2r)^d (1+e) c_1 r^{\xi_1} e^{-q} c_2 r^{\xi_2} e^{(vt-r)/\ell_2} e^{q\frac{\ell_1}{\ell_2}} \tag{A.54}$$

$$\leq \sum_q 17 \times 2^d c_1 c_2 r^{\xi_1+\xi_2+d} e^{(vt-r)/\ell_2} \tag{A.55}$$

$$\leq 17 \times 2^d c_1 c_2 r^{\xi_1+\xi_2+d+1} e^{(vt-r)/\ell_2}, \tag{A.56}$$

where we have used $\ell_1 \leq \ell_2$ and the fact that there are at most $\frac{r-v_1 t}{\ell_1} \leq r$ different q.

Combining Eqs. (A.50), (A.52) and (A.56) with $c_1, c_2 \geq 1$, $d \geq 1$ and $\xi_1, \xi_2 \geq 0$, we have

$$\|\mathbb{P}_r e^{\mathcal{L}_2 t} e^{\mathcal{L}_1 t} |O)\| \leq 2^{d+5} c_1 c_2 r^{\xi_1+\xi_2+d+1} e^{(vt-r)/\ell_2}, \tag{A.57}$$

with $v = v_1 + v_2$. Therefore, the lemma follows. □

A.1.1.2 Proof of Corollary 4

In this section, we prove Corollary 4, an application of Lemma 14 which extends the validity of a bound from $t \leq \Delta t$ to arbitrary time.

Proof. The lemma clearly holds for $k = 0$. So we will prove it by induction. Suppose Eq. (A.41) holds for some $k \in \mathbb{N}$. We will prove that it holds for $k + 1$.

The strategy is to apply the assumed bound for k [Eq. (A.41)] twice:

$$\|\mathbb{P}_r e^{\mathcal{L}t}|O)\| = \|\mathbb{P}_r e^{\mathcal{L}t/2} e^{\mathcal{L}t/2}|O)\|, \tag{A.58}$$

where the evolutions under $e^{\mathcal{L}t/2}$ can be bounded by the assumed bound because $t/2 \leq 2^k\Delta t$. We then use Lemma 14 to merge the two identical bounds with $v_1 = v_2 \to v/2$, $\ell_1 = \ell_2 \to \ell$, $c_1 = c_2 \to c_k$, $\xi_1 = \xi_2 \to \xi_k$:

$$\|\mathbb{P}_r e^{\mathcal{L}t}|O)\| \leq 2^{d+5} c_k^2 r^{2\xi_k+d+1} e^{(vt-r)/\ell}. \tag{A.59}$$

We choose

$$c_{k+1} = 2^{d+5}c_k^2 \qquad \Rightarrow c_k = 2^{(d+5)(2^k-1)}c_0^{2^k} \tag{A.60}$$

$$\xi_{k+1} = 2\xi_k + d + 1 \qquad \Rightarrow \xi_k = (2^k-1)(d+1) + 2^k\xi_0. \tag{A.61}$$

Therefore, by induction, Eq. (A.41) holds for $k+1$.

Next, to prove Eq. (A.42), we choose $k = \lceil \log_2(t/\Delta t)\rceil$ so that $t \leq 2^k\Delta t$. We also have $2^k \leq \frac{2t}{\Delta t}$. Therefore, $c_k \leq (2^{d+5}c_0)^{\frac{2t}{\Delta t}}$, $\xi_k \leq (d+1+\xi_0)\frac{2t}{\Delta t}$. Plugging them into Eq. (A.41), we have

$$\|\mathbb{P}_r e^{\mathbb{P}_r t}|O)\| \leq (2^{d+5}c_0)^{\frac{2t}{\Delta t}} r^{(d+1+\xi_0)\frac{2t}{\Delta t}} e^{\frac{vt-r}{\ell}} = e^{\chi\frac{t}{\Delta t}+\frac{vt-r}{\ell}}, \tag{A.62}$$

with $\chi = 2[\log(2^{d+5}c_0) + (d+1+\xi_0)\log r]$. □

A.1.2 Proof of Lemma 13

In this section, we prove Lemma 13.

Proof. First, we need the following lemma, which uses an existing bound to prove a tighter bound. We will use the lemma recursively to prove the nearly optimal bound in Lemma 13.

Lemma 15. *Let* $\eta \in (0, \frac{1}{\alpha - d})$ *be an arbitrary constant and*

$$\delta = \frac{2d+1}{(2d+1-\alpha)(1+\eta(2d+1-\alpha))} \tag{A.63}$$

be another constant. Suppose there exist constants $\gamma, C, c \geq 0$, $\kappa \geq \delta$, *and* $\beta > d$ *such that*

$$\|\mathbb{P}_r e^{\mathcal{L}t} |O)\| \leq C \log^{\kappa} r_* \frac{t^{\gamma}}{r^{\beta}}, \tag{A.64}$$

holds for all $t^{\gamma} \leq c r^{\beta} / \log^{\delta} r_*$. *Then, there exist constants* $C', c' > 0$, *and* $\kappa' > \delta$ *such that*

$$\|\mathbb{P}_r e^{\mathcal{L}t} |O)\| \leq C' \log^{\kappa'} r_* \frac{t^{\gamma'}}{r^{\beta'}}, \tag{A.65}$$

holds for all $t^{\gamma'} \leq c' r^{\beta'} / \log^{\delta} r_*$*, where*

$$\kappa' = \max\left\{\kappa - \frac{\delta(\beta - d)}{\beta} + \frac{\alpha - d}{2d + 1 - \alpha}, \delta\right\}, \tag{A.66}$$

$$\gamma' = \gamma d/\beta + 1 - \eta(\alpha - d), \tag{A.67}$$

$$\beta' = \alpha - d - \eta(\alpha - 2d)(\alpha - d) > d. \tag{A.68}$$

Proof. Let $V = H - H_n$ to be the sum over interactions of range more than ℓ_n. We have [34]

$$e^{\mathcal{L}t} |O) = e^{\mathcal{L}_n t} |O) + \sum_{h_{ij}} \int_0^t ds\, e^{\mathcal{L}(t-s)} \mathcal{L}_{h_{ij}} e^{\mathcal{L}_n s} |O)\,, \tag{A.69}$$

where the sum is over all h_{ij} in V. The first term is the evolution under H_n, which we can bound using Eq. (A.8). Our task is to bound the second term.

Without loss of generality, we assume $i \leq j$. Because $\mathcal{L}_{h_{ij}} |O)$ only acts nontrivially on the part of O supported at least a distance $\mathrm{dist}(i, 0)$ from the origin, we can insert $\mathbb{P}_{\mathrm{dist}(i,0)}$ in the middle of the intergrand and use the triangle inequality:

$$\|\mathbb{P}_r e^{\mathcal{L}t} |O)\| \leq \|\mathbb{P}_r e^{\mathcal{L}_n t} |O)\| + \|\mathbb{P}_r \sum_{h_{ij}} \int_0^t ds e^{\mathcal{L}(t-s)} \mathcal{L}_{h_{ij}} \mathbb{P}_{\mathrm{dist}(i,0)} e^{\mathcal{L}_n s} |O)\| \tag{A.70}$$

$$\leq \|\mathbb{P}_r e^{\mathcal{L}_n t} |O)\| + 4 \sum_{h_{ij}} \int_0^t ds \|h_{ij}\| \|\mathbb{P}_{\mathrm{dist}(i,0)} e^{\mathcal{L}_n s} |O)\|. \tag{A.71}$$

Because $\|h_{ij}\| \leq 1/\mathrm{dist}(i, j)^{\alpha}$ and $\mathrm{dist}(i, j) > \ell_n$, there exist a constant K_1 such

that $\sum_{j:\text{dist}(i,j)>\ell_n} \|h_{ij}\| \le K_1/\ell_n^{\alpha-d}$ for all $i \in \Lambda$. Therefore, we have

$$\|\mathbb{P}_r e^{\mathcal{L}t} |O)\| \le \|\mathbb{P}_r e^{\mathcal{L}_n t} |O)\| + \frac{4K_1}{\ell_n^{\alpha-d}} \int_0^t ds \sum_i \|\mathbb{P}_{\text{dist}(i,0)} e^{\mathcal{L}_n s} |O)\|, \tag{A.72}$$

We then consider two cases for the sum over i. If $\text{dist}(i,0)^\beta \le \frac{1}{c}\log^\delta(r_*)\ s^\gamma$, we use a trivial bound on the projection:

$$\sum_{i:\text{dist}(i,0)^\beta \le \frac{1}{c}\log^\delta(r_*)\ s^\gamma} \|\mathbb{P}_{\text{dist}(i,0)} e^{\mathcal{L}_n s} |O)\| \le (2c^{-1/\beta}(\log r_*)^{\delta/\beta} s^{\gamma/\beta})^d \times 2 \tag{A.73}$$

$$= 2^{d+1} c^{-d/\beta} (\log r_*)^{\delta d/\beta} s^{\gamma d/\beta}. \tag{A.74}$$

Otherwise, if $\text{dist}(i,0)^\beta > \frac{1}{c}\log^\delta(r_*)\ s^\gamma$, we apply Eq. (A.64):

$$\sum_{i:\text{dist}(i,0)^\beta > \frac{1}{c}\log^\delta(r_*)\ s^\gamma} \|\mathbb{P}_{\text{dist}(i,0)} e^{\mathcal{L}_n s} |O)\| \le C \log^\kappa r_* \sum_{i:\text{dist}(i,0)^\beta > \frac{1}{c}\log^\delta(r_*)\ s^\gamma} \frac{s^\gamma}{\text{dist}(i,0)^\beta} \tag{A.75}$$

$$\le C K_2 \log^\kappa r_* \frac{s^\gamma}{[c^{-1/\beta}(\log r_*)^{\delta/\beta} s^{\gamma/\beta}]^{\beta-d}} \tag{A.76}$$

$$\le C K_2 c^{\frac{\beta-d}{\beta}} (\log r_*)^{\kappa - \frac{\delta(\beta-d)}{\beta}}\ s^{\gamma d/\beta}, \tag{A.77}$$

where K_2 is a constant such that

$$\sum_{i:\text{dist}(i,0)>a} \frac{1}{\text{dist}(i,0)^\beta} \le \frac{K_2}{a^{\beta-d}}, \tag{A.78}$$

for all $a > 0$. Such a constant K_2 exists because $\beta > d$ by assumption.

Combining Eqs. (A.74) and (A.77) and accounting for $\kappa \geq \delta$, we can upper bound

$$\frac{4K_1}{\ell_n^{\alpha-d}} \int_0^t ds \sum_i \|\mathbb{P}_{\text{dist}(i,0)} e^{\mathcal{L}_n s} |O)\| \leq K (\log r_*)^{\kappa - \frac{\delta(\beta-d)}{\beta}} \frac{t^{\frac{\gamma d}{\beta}+1}}{\ell_n^{\alpha-d}}, \tag{A.79}$$

where we absorb all constants into $K = 4K_1(2^{d+1}c^{-d/\beta} + CK_2/c^{\frac{\beta-d}{\beta}})\frac{\beta}{\gamma d+\beta}$. Substituting Eq. (A.79) in Eq. (A.72), we have a bound for the evolution under H:

$$\|\mathbb{P}_r e^{\mathcal{L}t} |O)\| \leq e^{\frac{v_n t - r}{\ell_n}} + K (\log r_*)^{\kappa - \frac{\delta(\beta-d)}{\beta}} \frac{t^{\frac{\gamma d}{\beta}+1}}{\ell_n^{\alpha-d}} \tag{A.80}$$

We now substitute the values of v_n and ℓ_n into the bound. Recall from Eq. (A.3) that

$$n = \left\lfloor \frac{1}{\log L} \log \left[r \left(\frac{t}{r^{\alpha-2d}} \right)^{\eta} \right] \right\rfloor, \tag{A.81}$$

where $\eta \in (0, \frac{1}{\alpha-d})$ is an arbitrary small constant. With this choice, we can bound ℓ_n from both above and below:

$$r \geq r \left(\frac{t}{r^{\alpha-2d}} \right)^{\eta} \geq \ell_n = L^n \geq \frac{r}{L} \left(\frac{t}{r^{\alpha-2d}} \right)^{\eta} = \frac{r}{(\xi \log r_*)^{1/(2d+1-\alpha)}} \left(\frac{t}{r^{\alpha-2d}} \right)^{\eta}. \tag{A.82}$$

With $v_1 = 4e\tau\ell_1$, $x = \xi \log r_* = L^{2d+1-\alpha}$, we also have a bound for v_n from Eq. (A.10):

$$v_n \leq r^{2d+1-\alpha} \left(\frac{t}{r^{\alpha-2d}} \right)^{\eta(2d+1-\alpha)} \frac{4e\tau x^{1/(2d+1-\alpha)} + \nu\lambda(n-1)x^{(2d+1)/(2d+1-\alpha)}}{x}. \tag{A.83}$$

Assuming that $r_* \geq e^{e^{2d+1-\alpha}/\xi}$ so that $\log L \geq 1$ and $x \geq 1$, we have $n \leq \log r / \log L \leq \log r_* = x/\xi$. We can then crudely upper bound

$$v_n \leq (4e\tau + \nu\lambda/\xi)(\xi \log r_*)^{\frac{2d+1}{2d+1-\alpha}} \frac{r}{t} \left(\frac{t}{r^{\alpha-2d}} \right)^{1+\eta(2d+1-\alpha)} \tag{A.84}$$

$$= K_3 (\log r_*)^{\frac{2d+1}{2d+1-\alpha}} \frac{r}{t} \left(\frac{t}{r^{\alpha-2d}} \right)^{1+\eta(2d+1-\alpha)} . \tag{A.85}$$

where K_3 is a constant. Assuming

$$t \leq r^{\alpha-2d} / [2K_3 (\log r_*)^{\frac{2d+1}{2d+1-\alpha}}]^{\frac{1}{1+\eta(2d+1-\alpha)}} \tag{A.86}$$

so that $v_n t \leq r/2$, we can simplify the first term of Eq. (A.80):

$$e^{\frac{v_n t - r}{\ell_n}} \leq e^{-\frac{r}{2\ell_n}} \leq \exp\left[-\frac{1}{2} \left(\frac{r^{\alpha-2d}}{t} \right)^{\eta} \right]. \tag{A.87}$$

Similarly, the second term of Eq. (A.80) can be simplified to

$$K(\log r_*)^{\kappa - \frac{\delta(\beta-d)}{\beta}} \frac{t^{\gamma d/\beta + 1}}{\ell_n^{\alpha-d}} \leq K(\log r_*)^{\kappa - \frac{\delta(\beta-d)}{\beta}} t^{\gamma d/\beta+1} \left[\frac{r}{(\xi \log r_*)^{1/(2d+1-\alpha)}} \left(\frac{t}{r^{\alpha-2d}} \right)^{\eta} \right]^{d-\alpha} \tag{A.88}$$

$$= K \xi^{(\alpha-d)/(2d+1-\alpha)} (\log r_*)^{\kappa - \frac{\delta(\beta-d)}{\beta} + \frac{\alpha-d}{2d+1-\alpha}} \frac{t^{\gamma d/\beta + 1 - \eta(\alpha-d)}}{r^{\alpha-d-\eta(\alpha-2d)(\alpha-d)}} \tag{A.89}$$

$$= K_4 \log^{\kappa'} r_* \frac{t^{\gamma'}}{r^{\beta'}}, \tag{A.90}$$

where K_4, κ', γ' are constants. In particular, $\gamma' = \gamma d/\beta + 1 - \eta(\alpha - d)$ and $\beta' =$

$\alpha - d - \eta(\alpha - 2d)(\alpha - d) > d$. Combining Eqs. (A.87) and (A.90), we get a bound for the evolution under H.

We now simplify the bound by considering t such that

$$t^{\gamma'} \leq \frac{c' r^{\beta'}}{(\log r_*)^{\delta\gamma'}} \tag{A.91}$$

for some constant c' and $\delta = \frac{2d+1}{(2d+1-\alpha)(1+\eta(2d+1-\alpha))}$. Since $\beta/\gamma \leq \alpha - 2d$ by assumption, we also have

$$\frac{\beta'}{\gamma'} = \frac{\alpha - d - \eta(\alpha - 2d)(\alpha - d)}{\frac{\gamma d}{\beta} + 1 - \eta(\alpha - d)} \leq \frac{\alpha - d - \eta(\alpha - 2d)(\alpha - d)}{\frac{d}{\alpha - 2d} + 1 - \eta(\alpha - d)} = \alpha - 2d. \tag{A.92}$$

Therefore, with $c' = (2K_3)^{\frac{-\gamma'}{(1+\eta(2d+1-\alpha))}}$, Eq. (A.91) satisfies the condition in Eq. (A.86). In addition, for $r^{\alpha-2d} \geq t$, there exists a constant K_5 such that

$$\exp\left[-\frac{1}{2}\left(\frac{r^{\alpha-2d}}{t}\right)^{\eta}\right] \leq K_5 \left(\frac{t}{r^{\alpha-2d}}\right)^{\gamma'} \leq K_5 \frac{t^{\gamma'}}{r^{\beta'}}, \tag{A.93}$$

where we have again used $\gamma'(\alpha - 2d) \geq \beta'$ in the last inequality. Replacing Eq. (A.87) by Eq. (A.93) and combining with Eq. (A.90), we arrive at a bound

$$\|\mathbb{P}_r e^{\mathcal{L}t} |O)\| \leq C' \log^{\kappa'} r_* \frac{t^{\gamma'}}{r^{\beta'}}, \tag{A.94}$$

for all $t^{\gamma'} \leq c' r^{\beta'} / \log^{\delta} r_*$, where $C' \geq K_4 + K_5$ and c' are constants,

$$\kappa' = \kappa - \frac{\delta(\beta - d)}{\beta} + \frac{\alpha - d}{2d + 1 - \alpha}, \tag{A.95}$$

$$\gamma' = \gamma d/\beta + 1 - \eta(\alpha - d), \tag{A.96}$$

$$\beta' = \alpha - d - \eta(\alpha - 2d)(\alpha - d) > d. \tag{A.97}$$

If $\kappa' < \delta$, we simply replace κ' by δ in Eq. (A.94). Such replacement can only increase the bound in Eq. (A.94). Therefore, Lemma 15 follows. □

We now use Lemma 15 to prove Lemma 13. To satisfy the assumption of Lemma 15, we start with the bound in Ref. [19]: There exist constants K_6, K_7, and v_F such that

$$\| [O', e^{\mathcal{L}t} O] \| \leq K_6 \exp\left(v_F t - \frac{r}{t^{(1+d)/(\alpha-2d)}} \right) + K_7 \frac{t^{\frac{\alpha(\alpha-d+1)}{\alpha-2d}}}{r^{\alpha}}, \tag{A.98}$$

for all single-site, unit-norm operators O' supported a distance r from O. We consider the regime

$$t^{\frac{\alpha(\alpha-d+1)}{\alpha-2d}} \leq c r^{\alpha-d} / \log^{\delta} r_* \leq c r^{\alpha-d} \leq c r^{\alpha}, \tag{A.99}$$

where we choose $c = (2v_F)^{-\alpha}$ so that

$$v_F t \leq \frac{r}{2t^{(1+d)/(\alpha-2d)}}. \tag{A.100}$$

Therefore, there exists a constant K_8 such that

$$K_6 \exp\left(v_F t - \frac{r}{t^{(1+d)/(\alpha-2d)}}\right) \leq K_6 \exp\left(-\frac{r}{2t^{(1+d)/(\alpha-2d)}}\right) \tag{A.101}$$

$$\leq K_8 \left(\frac{t^{(1+d)/(\alpha-2d)}}{r}\right)^{\frac{\alpha(\alpha-d+1)}{1+d}} \leq K_8 \frac{t^{\alpha(\alpha-d+1)/(\alpha-2d)}}{r^\alpha} \tag{A.102}$$

holds for all t satisfying Eq. (A.99). In the last inequality, we have used $\alpha - d + 1 \geq d + 1$ to lower bound the exponent of r. Substituting Eq. (A.102) into Eq. (A.98), we get a simplified version of the bound in Ref. [19]:

$$\| \left[O', e^{\mathcal{L}t} O\right] \| \leq K_9 \frac{t^{\frac{\alpha(\alpha-d+1)}{\alpha-2d}}}{r^\alpha}, \tag{A.103}$$

where $K_9 = K_7 + K_8$. Applying Lemma 4 in Ref. [25], there exists a constant K_{10} such that:

$$\| \mathbb{P}_r e^{\mathcal{L}t} |O) \| \leq K_{10} \frac{t^{\frac{\alpha(\alpha-d+1)}{\alpha-2d}}}{r^{\alpha-d}} < K_{10} \log^\delta r_* \frac{t^{\frac{\alpha(\alpha-d+1)}{\alpha-2d}}}{r^{\alpha-d}}, \tag{A.104}$$

where the additional factor $-d$ in the exponent of r comes from "integrating" over sites that are at least a distance r from the origin. Equation (A.104) satisfies the assumption of Lemma 15, with $C \to K_{10}, c \to (2v_F)^{-\alpha}, \kappa \to \delta, \gamma \to \frac{\alpha(\alpha-d+1)}{\alpha-2d}, \beta \to \alpha - d$. Therefore, by the lemma, there exist constants C_1, c_1, κ_1 such that

$$\| \mathbb{P}_r e^{\mathcal{L}_r} |O) \| \leq C_1 \log^{\kappa_1} r_* \frac{t^{\gamma_1}}{r^{\beta_1}} \tag{A.105}$$

holds for all $t^{\gamma_1} \leq c_1 r^{\beta_1} / \log^{\delta} r_*$, where

$$\gamma_1 = \frac{\alpha(\alpha - d + 1)}{\alpha - 2d} \frac{d}{\alpha - d} + 1 - \eta(\alpha - d), \tag{A.106}$$

$$\beta_1 = \alpha - d - \eta(\alpha - 2d)(\alpha - d). \tag{A.107}$$

Equation (A.105) again satisfies the assumption of Lemma 15. Applying the lemma again with $\gamma \to \gamma_1, \beta \to \beta_1$, we obtain

$$\|\mathbb{P}_r e^{\mathcal{L}_r} |O)\| \leq C_2 \log^{\kappa_2} r_* \frac{t^{\gamma_2}}{r^{\beta_2}} \tag{A.108}$$

for some constants $C_2, \kappa_2, \beta_2 = \beta_1$, and

$$\gamma_2 = \frac{\gamma_1 d}{\beta_1} + 1 - \eta(\alpha - d) \equiv f(\gamma_1). \tag{A.109}$$

After applying Lemma 15 for m times, we obtain

$$\|\mathbb{P}_r e^{\mathcal{L}_r} |O)\| \leq C_m \log^{\kappa_m} r_* \frac{t^{\gamma_m}}{r^{\beta_m}} = C_m \log^{\kappa_m} r_* \left(\frac{t}{r^{\beta_m/\gamma_m}} \right)^{\gamma_m} \tag{A.110}$$

for some constants $C_m, \kappa_m, \beta_m = \beta_1$, and

$$\gamma_m = f^{\circ(m-1)}(\gamma_1), \tag{A.111}$$

where $f^{\circ(m-1)}$ denotes the $(m-1)$-th composition of the function f. It is straightforward

to show that

$$\lim_{\eta\to 0}\lim_{m\to\infty}\gamma_m = \lim_{\eta\to 0}\frac{\alpha-d-\eta(\alpha-2d)(\alpha-d)}{\alpha-2d} = \frac{\alpha-d}{\alpha-2d}, \tag{A.112}$$

$$\lim_{\eta\to 0}\lim_{m\to\infty}\frac{\beta_m}{\gamma_m} = \alpha-2d. \tag{A.113}$$

Therefore, for all $\varepsilon>0$, there exist $m\geq 1, \eta\in(0,\frac{1}{\alpha-d})$ such that $\beta_m/\gamma_m \geq \alpha-2d-\varepsilon$ and $\gamma_m \geq \frac{\alpha-d}{\alpha-2d}-\varepsilon$. We obtain

$$\|\mathbb{P}_r e^{\mathcal{L}_r}|O)\| \leq C_m \log^{\kappa_m} r_* \left(\frac{t}{r^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\varepsilon}, \tag{A.114}$$

which holds for all $t \leq c_m^{1/\gamma_m} r^{\alpha-2d-\varepsilon}/(\log r_*)^{\delta/\gamma_m} \leq r^{\alpha-2d-\varepsilon}$. Lemma 13 thus follows.

□

A.1.3 Removing the Dependence on the Lattice Size

In this section, we use Lemma 13 to prove Theorem 11 by removing the dependence on r_*.

Proof. Let $H_{\text{out}} = \mathbb{P}_{r_0} H$ denote the terms of the Hamiltonian H that have support outside a distance r_0 from the origin, $H_{\text{in}} = H - H_{\text{out}}$ be the rest of the Hamiltonian, and $\mathcal{L}_{\text{out}}$, $\mathcal{L}_{\text{in}}$ are the corresponding Liouvillians. Using the triangle inequality, we have

$$\|\mathbb{P}_r e^{\mathcal{L}t}|O)\| \leq \|\mathbb{P}_r e^{\mathcal{L}_{\text{in}}t}|O)\| + \sum_{h_{ij}}\|\mathbb{P}_r\int_0^t ds\, e^{\mathcal{L}(t-s)}\mathcal{L}_{h_{ij}} e^{\mathcal{L}_{\text{in}}s}|O)\|, \tag{A.115}$$

where the sum is taken over terms h_{ij} in H_{out}. Without loss of generality, we assume

$\text{dist}(i,0) \leq \text{dist}(j,0)$, which implies $\text{dist}(j,0) \geq r_0$. In addition, since $e^{\mathcal{L}_{\text{in}} s}|O)$ is supported entirely within the radius r_0 from the origin, only terms where $\text{dist}(i,0) \leq r_0$ contribute to the above sum. We consider two cases: $\text{dist}(i,0) > r_0/2$ and $\text{dist}(i,0) \leq r_0/2$.

In the former case, we insert $\mathbb{P}_{\text{dist}(i,0)}$ in the middle of the integrand and bound

$$\sum_{h_{ij}:\text{dist}(i,0)\in(\frac{r_0}{2},r_0]} \|\mathbb{P}_r \int_0^t ds\; e^{\mathcal{L}(t-s)} \mathcal{L}_{h_{ij}} \mathbb{P}_{\text{dist}(i,0)} e^{\mathcal{L}_{\text{in}} s}|O)\|$$

$$\leq 4 \sum_{h_{ij}:\text{dist}(i,0)\in(\frac{r_0}{2},r_0]} \|h_{ij}\| \int_0^t ds \|\mathbb{P}_{\text{dist}(i,0)} e^{\mathcal{L}_{\text{in}} s}|O)\| \tag{A.116}$$

$$\leq 4K_1 C \log^{\kappa}(2r_0) \sum_{i:\text{dist}(i,0)\in(\frac{r_0}{2},r_0]} \int_0^t ds \left(\frac{t}{\text{dist}(i,0)^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\varepsilon} \tag{A.117}$$

$$\leq K_2 r_0^d t \left(\frac{t}{r_0^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\varepsilon}, \tag{A.118}$$

where K_2 is a constant. We have used Lemma 13 to bound the evolution under $e^{\mathcal{L}_{\text{in}} s}$, which is supported entirely within a truncated lattice of diameter $2r_0$, and used the fact that the interaction h_{ij} decays as a power law with an exponent $\alpha > 2d$ to bound the sum over j by a constant. We require $t \leq c_1 r^{\alpha-2d-\varepsilon}/\log^{\delta}(2r_0)$, for some constant c_1, δ, to satisfy the conditions of Lemma 13.

On the other hand, when $\text{dist}(i,0) \leq r_0/2$, we have $\text{dist}(i,j) \geq r_0/2$. Therefore, there exists a constant c_2 such that $\sum_j \|h_{ij}\| \leq c_2/r_0^{\alpha-d}$ for all i. We can then bound

$$\sum_{h_{ij}:\text{dist}(i,0)\leq\frac{r_0}{2}} \|\mathbb{P}_r \int_0^t ds\; e^{\mathcal{L}(t-s)} \mathcal{L}_{h_{ij}} e^{\mathcal{L}_{\text{in}} s}|O)\| \leq 4 \sum_{i:\text{dist}(i,0)\leq\frac{r_0}{2}} \frac{c_2}{r_0^{\alpha-d}} \int_0^t ds \leq K_3 \frac{t}{r_0^{\alpha-2d}}, \tag{A.119}$$

for some constant K_3.

Using Lemma 13 on the first term of Eq. (A.115) and combining with Eqs. (A.118) and (A.119), we have:

$$\|\mathbb{P}_r e^{\mathcal{L}t}\,|O)\| \leq C\log^{\kappa}(2r_0)\left(\frac{t}{r^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\varepsilon} + K_2 r_0^d t\left(\frac{t}{r_0^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\varepsilon} + K_3\frac{t}{r_0^{\alpha-2d}}. \tag{A.120}$$

We choose $r_0 = r^{\xi}$, where

$$\xi = \left(1+\frac{\alpha-d}{\alpha-2d}-\varepsilon\right)\frac{\alpha-2d-\varepsilon}{\alpha-2d-\varepsilon\frac{(\alpha-2d)^2+\alpha-d}{\alpha-2d}+\varepsilon^2} \geq \frac{\alpha-d}{\alpha-2d}, \tag{A.121}$$

and we require $\varepsilon \leq (\alpha-2d)^2/[(\alpha-2d)^2+\alpha-d]$ so that the lower bound on ξ holds. Under this choice,

$$K_2 r_0^d t\left(\frac{t}{r_0^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\varepsilon} = K_2\left(\frac{t}{r^{\alpha-2d-\varepsilon}}\right)^{1+\frac{\alpha-d}{\alpha-2d}-\varepsilon} \leq K_4\left(\frac{t}{r^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\varepsilon}, \tag{A.122}$$

for all $t \leq c_1 r^{\alpha-2d-\varepsilon}$, where K_4 is a constant. In addition, for $\varepsilon \leq (\alpha-2d)^2/[(\alpha-2d)^2+\alpha-d]$, $\xi \geq (\alpha-d)/(\alpha-2d)$ and, therefore,

$$K_3\frac{t}{r_0^{\alpha-2d}} \leq K_3\frac{t}{r^{\alpha-d}}. \tag{A.123}$$

Combining Eqs. (A.121)–(A.123), we have

$$\|\mathbb{P}_r e^{\mathcal{L}t}|O)\| \leq K_5 \log^\kappa(r) \left(\frac{t}{r^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\varepsilon} + K_6 \frac{t}{r^{\alpha-d}}, \tag{A.124}$$

which holds for all $t \leq c_1 r^{\alpha-2d-\varepsilon}/\log^\delta(2r^\xi)$ for some constants K_5, K_6 independent of t, r.

Next, we simplify Eq. (A.124) by "hiding" the factor $\log^\kappa r$ inside the constant ε. Specifically, there exist a constant K_7 such that $\log^\kappa r \leq K_7 r^{\varepsilon'}$, where $\varepsilon' = \frac{\varepsilon}{2}\left(\frac{\alpha-d}{\alpha-2d} - \frac{\varepsilon}{2}\right)$, and constants K_8, K_9 such that

$$\|\mathbb{P}_r e^{\mathcal{L}t}|O)\| \leq K_8 \log^\kappa(r) \left(\frac{t}{r^{\alpha-2d-\varepsilon/2}}\right)^{\frac{\alpha-d}{\alpha-2d}-\frac{\varepsilon}{2}} + K_9 \frac{t}{r^{\alpha-d}} \tag{A.125}$$

$$\leq K_7 K_8 \left(\frac{t}{r^{\alpha-2d-\varepsilon}}\right)^{\frac{\alpha-d}{\alpha-2d}-\frac{\varepsilon}{2}} + K_9 \frac{t}{r^{\alpha-d}}, \tag{A.126}$$

which holds for all $t \leq c_1 r^{\alpha-2d-\varepsilon/2}/\log^\delta(2r^\xi)$. In addition, there exists a constant K_{10} such that $K_{10}\log^\delta(2r^\xi) \leq r^{\varepsilon/2}$ for all $r \geq 1$. By requiring that $t \leq K_{10} c_1 r^{\alpha-2d-\varepsilon}$, we also ensure $t \leq c_1 r^{\alpha-2d-\varepsilon/2}/\log^\delta(2r^\xi)$. Therefore, Theorem 11 follows with $c \to c_1 K_{10}, C_1 \to K_7 K_8$, and $C_2 \to K_9$.

□

A.2 Applications of Theorem 11

We discussed in the main text that the tightened light cone and nearly optimal tail in Theorem 11 improved the scaling for various applications of Lieb-Robinson bounds to problems of physical interest in the regime $2d < \alpha < 2d + 1$ Here we provide some

mathematical details to justify those assertions. We also provide a table briefly summarizing the bounds we will use to compare, where we consider each bound to take the form $\|[A(t), B]\| \leq ct^\gamma / r^\beta$ for some constants c, γ and β, where A is single-site, but B may generally be some large multi-site operator.

Bound	Light cone	Tail	γ	β	γ'	β'
This work (B1)	$t \gtrsim r^{\alpha-2d}$	$1/r^{\alpha-d}$	$\frac{\alpha-d}{\alpha-2d}$	$\alpha-d$	$\gamma+1$	β
Ref. [19] (B2)	$t \gtrsim r^{\frac{(\alpha-d)(\alpha-2d)}{\alpha(\alpha-d+1)}}$	$1/r^{\alpha-d}$	$\frac{\alpha(\alpha-d+1)}{\alpha-2d}$	$\alpha-d$	$\gamma+1$	β
Ref. [21] (B3)	$t \gtrsim r^{\frac{\alpha-2d}{\alpha-d}}$	$1/r^{\alpha-2d}$	$\alpha-d$	$\alpha-2d$	γ	β

Table A.1: Comparison of Lieb-Robinson bounds for $2d < \alpha < 2d+1$. We name the bounds B1, B2, and B3 for brevity. We ignore the arbitrarily small parameter ε in B1 for simplicity, as it does not affect the conclusions.

We first consider the application of the bound on the growth of connected correlators. Consider two unit-norm, single-site observables A and B initially supported on sites x and y, respectively, such that x and y are separated by a distance r. Let $|\psi\rangle$ be a product state between $\mathcal{B}_{r/2}(x)$ and $\mathcal{B}_{r/2}(y)$, where $\mathcal{B}_{r/2}(x)$ is the ball of radius $r/2$ around x. The connected correlator is defined by

$$C(r,t) \equiv \langle A(t)B(t)\rangle - \langle A(t)\rangle \langle B(t)\rangle , \tag{A.127}$$

where $\langle\cdot\rangle \equiv \langle\psi|\cdot|\psi\rangle$. Define $\tilde{A}(t) \equiv \mathrm{Tr}_{\mathcal{B}^c_{r/2}(x)}[A(t)]$ and $\tilde{B}$ similarly. It is elementary to bound $C(r,t)$ by

$$C(r,t) \leq 2\|A(t) - \tilde{A}(t)\| + 2\|B(t) - \tilde{B}(t)\|. \tag{A.128}$$

That is, the connected correlator is controlled by the error in truncating $A(t)$ and $B(t)$ to

within a ball of radius $r/2$ around their initial support. A simple result from Ref. [36] allows us to bound this error

$$\|A(t) - \tilde{A}(t)\| \leq \int_{\mathcal{B}^c_{r/2}(x)} dU \|[U, A(t)]\| \leq c \frac{t^\gamma}{(r/2)^\beta}, \tag{A.129}$$

where dU is the Haar measure on unitaries supported outside a ball of radius $r/2$ around x. Thus, for a given Lieb-Robinson bound

$$C(r,t) \leq 2^{\beta+2} c \frac{t^\gamma}{r^\beta}. \tag{A.130}$$

Ignoring constants and focusing on the asymptotics with respect to t and r, we see that

$$R_{12} \equiv \frac{\text{B1}}{\text{B2}} \sim \left(\frac{t^{\alpha-d}}{t^{\alpha(\alpha-d+1)}} \right)^{\frac{1}{\alpha-2d}}, \tag{A.131}$$

$$R_{13} \equiv \frac{\text{B1}}{\text{B3}} \sim \left(\frac{t^{\frac{1}{\alpha-2d}}}{t} \right)^{\alpha-d} r^{-d}. \tag{A.132}$$

Thus, as t increases, the tighter light cone of B1 leads to significant improvement in bounding the connected correlator as compared to B2. While B1 has a slightly worse time-dependence than B3 (as $0 < \alpha - 2d < 1$), it has a much better r-dependence. And, of course, when taken together, B1 follows a tighter light cone than B3, leading to an overall more useful bound. Thus, while B3 may strictly have a better time-dependence, B1 provides the tightest holistic bound on the growth of connected correlators.

A nearly identical calculation allows us to place stricter bounds on the time required to generate topologically ordered states from topologically trivial ones. We define topo-

logically ordered states as follows: consider a lattice Λ with diameter L and $O(L^d)$ sites. We say that a set of orthonormal states $\{|\psi_1\rangle, \ldots, |\psi_k\rangle\}$ are topologically ordered if there exists a constant δ such that

$$\varepsilon \equiv \sup_O \max_{1 \leq i,j \leq k} \{|\langle\psi_i| O |\psi_j\rangle - \langle\psi_j| O |\psi_i\rangle, 2\langle\psi_i| O |\psi_j\rangle|\} \tag{A.133}$$

is bounded $\varepsilon = O(L^{-\delta})$. The supremum is taken over operators O supported on a subset of the lattice with diameter $\ell' < L$, so ε essentially measures the ability to distinguish between states $|\psi_i\rangle$ using an operator O supported on only a fraction of the lattice. In contrast, we say the states are topologically trivial if ε is independent of L. Given a set of topologically trivial states $\{|\phi_i\rangle\}_i$ and a set of topologically ordered states $\{|\psi_i\rangle\}_i$, the question is how long it takes to generate a unitary U such that $U|\phi_i\rangle = |\psi_i\rangle$ for all i using a power-law Hamiltonian. Ref. [25] proves that this time is controlled by the time it takes $\|O(t) - O(t, \ell')\|$ to become non-vanishing in L, where $O(t, \ell')$ is the truncation of the time evolution of O to a radius ℓ'. This expression is bounded in the exact same way as the connected correlator was, and so we see the same improvement from B1 as compared to both B2 and B3.

Finally, we consider the task of simulating the evolution of a local observable under a power-law Hamiltonian H using quantum simulation algorithms. In contrast to the earlier applications, it is not sufficient to simply truncate the time-evolved observable to the light cone. Instead, to simulate the observable, we need to construct the Hamiltonian that generates the dynamics of the observable inside the light cone.

Let A be a unit-norm, single-site observable originally supported on site x, and

consider $A(t)$ its evolution under a 2-local power-law Hamiltonian H. Define H_r to be the Hamiltonian constructed by taking terms of H that are fully supported within $\mathcal{B}_r(x)$, and let $\tilde{A}(t)$ be $A(0)$ evolved under H_r (note that this is different than our previous definition of $\tilde{A}$). The question is how large r must be (i.e., how many terms of H must we simulate) for $\|A(t) - \tilde{A}(t)\|$ to have small error. Intuitively, this observable should be constrained to lie mostly within the light cone of a Lieb-Robinson bound for H as long as the tail of the bound decays sufficiently quickly, so we expect r to be related to the lightcone of our bounds. Refs. [21, 25] make this intuition rigorous and yield

$$\|A(t) - \tilde{A}(t)\| \lesssim \frac{t^{\gamma'}}{r^{\beta'}}, \tag{A.134}$$

where γ' and β' are listed in Table A.1. In particular, in order to ensure only a constant error, we must choose $r \sim t^{\gamma'/\beta'}$, which corresponds to simulating about $r^2 \sim t^{2d\gamma'/\beta'}$ terms of the Hamiltonian. We can compare this exponent $\phi \equiv \gamma'/\beta'$ between bounds:

$$\phi_{\text{B1}} - \phi_{\text{B2}} = -\frac{(\alpha-1)(\alpha-d)+\alpha}{(\alpha-d)(\alpha-2d)}, \tag{A.135}$$

$$\phi_{\text{B1}} - \phi_{\text{B3}} = -\frac{(\alpha-d)^2+d}{(\alpha-2d)(\alpha-d)}. \tag{A.136}$$

These differences are all negative for $2d < \alpha < 2d+1$, meaning the current work provides the tightest bound on how many terms must be kept to get constant error when simulating the evolution of local observables in this regime.

Appendix B: Supplemental Material for Chapter 5

B.1 Evaluations of the Sum in Lemma 3

In this section, we shall show how we bound δ_{trunc} from Eq. (5.10) (Appendix B.1.1) and δ_{overlap} from Eq. (5.11) (Appendix B.1.2) in the proof of Lemma 3.

B.1.1 Evaluation of δ_{trunc}

In this subsection, we provide explicit calculations of δ_{trunc} in Eq. (5.10). Recall that $\ell = \text{dist}\,(A, C)$ is the shortest distance between any two points in A and C. Therefore, $\|\vec{a} - \vec{c}\|$ is always greater than ℓ. For each $\vec{a} \in A$, let $\ell_{\vec{a}} = \text{dist}\,(\vec{a}, C)$ be the minimum distance from $\vec{a}$ to the set C and $C_{\vec{a}} = \{\vec{\imath} \in \Lambda : \text{dist}\,(\vec{a}, \vec{\imath}) \geq \ell_{\vec{a}}\}$. Clearly, C is a subset of $C_{\vec{a}}$. Therefore,

$$\|\delta_{\text{trunc}}\| = \|H_{A:C}\| \leq \sum_{\vec{a}\in A} \sum_{\vec{c}\in C} \frac{1}{\|\vec{a} - \vec{c}\|^{\alpha}} \tag{B.1}$$

$$\leq \sum_{\vec{a}\in A} \sum_{\vec{c}\in C_{\vec{a}}} \frac{1}{\|\vec{a} - \vec{c}\|^{\alpha}} = \sum_{\vec{a}\in A} \sum_{\substack{\vec{r} \\ \|\vec{r}\|\geq \ell_{\vec{a}}}} \frac{1}{\|\vec{r}\|^{\alpha}} \tag{B.2}$$

$$\leq \lambda_1 \sum_{\vec{a}\in A} \frac{1}{(\ell_{\vec{a}} - \sqrt{D})^{\alpha - D}}, \tag{B.3}$$

where λ_1 is a constant independent of $\vec{a}$ and the sum over $\vec{r}$ is bounded using Lemma 18 in Appendix B.7.

Next, to evaluate the sum over $\vec{a}$, we parameterize the sites in the set A by their distance to its boundary ∂A. Note that by assumption the interior of A is non-empty, so that $A \neq \partial A$. Roughly speaking, there will be at most $\mathcal{O}(\Phi(A))$ sites whose distances to the boundary ∂A is between ℓ and $\ell + \mu$, for each $\mu = 0, 1, \ldots$, where $\Phi(A)$ is the boundary area of A. Therefore, we have (see Lemma 22 in Appendix B.7.2 for a rigorous proof)

$$\|\delta_{\text{trunc}}\| \leq 2\eta\lambda_1\Phi(A)\sum_{\mu=0}^{\infty}\frac{1}{(\ell+\mu-\sqrt{D})^{\alpha-D}} \tag{B.4}$$

$$\leq 2\eta\lambda_1\lambda_2 2^{\alpha-D}\frac{\Phi(A)}{\ell^{\alpha-D-1}} = c_{\text{tr}}2^{\alpha}\frac{\Phi(A)}{\ell^{\alpha-D-1}}, \tag{B.5}$$

for $\ell > 2\sqrt{D}$, where λ_2 is a constant that arises after using Lemma 18 to bound the sum, and the factor $2^{\alpha-D}$ is because we lower bound $\ell - \sqrt{D} \geq \ell/2$ to simplify the expression. The constants are later absorbed into the definition of c_{tr}.

B.1.2 Evaluation of δ_{overlap}

In this section, we show how we bound δ_{overlap} from Eq. (5.11) in the proof of Lemma 3. To estimate δ_{overlap}, we use the following lemma, which generalizes a similar lemma in Ref. [76] to arbitrary, time-dependent Hamiltonians.

Lemma 16. *Let $\Omega \subset \Lambda$ be a subset of sites. Let $H_\Omega(t) = \sum_{i,j\in\Omega} h_{ij}(t)$ be the terms of $H_\Lambda(t)$ supported entirely on Ω. Let $O_X(\tau)$ be an observable supported on a subset X at*

a fixed time τ. We have:

$$\begin{aligned}
&\left\| \left(U_{0,t}^{\Lambda}\right)^{\dagger} O_X(\tau) U_{0,t}^{\Lambda} - \left(U_{0,t}^{\Omega}\right)^{\dagger} O_X(\tau) U_{0,t}^{\Omega} \right\| \\
&\leq \int_0^t ds \, \left\| \left[\left(U_{s,t}^{\Omega}\right)^{\dagger} O_X(\tau) U_{s,t}^{\Omega}, H_{\Lambda}(s) - H_{\Omega}(s) \right] \right\|, \qquad \text{(B.6)}
\end{aligned}$$

where $U_{0,t}^{\Lambda} = \mathcal{T} \exp\left(-i \int_0^t H_{\Lambda}(s) ds\right)$.

Proof. To prove the lemma, we shall move into the interaction picture of $H_{\Omega}(t)$ and treat $V(t) \equiv H_{\Lambda}(t) - H_{\Omega}(t)$ as a perturbation. Let $V_I(t) = \left(U_{0,t}^{\Omega}\right)^{\dagger} V(t) U_{0,t}^{\Omega}$ and $U_I(t) = \mathcal{T} \exp\left(-i \int_0^t V_I(s) ds\right)$ be respectively the Hamiltonian and the evolution operator in the interaction picture. We have:

$$\begin{aligned}
&\left\| \left(U_{0,t}^{\Lambda}\right)^{\dagger} O_X(\tau) U_{0,t}^{\Lambda} - \left(U_{0,t}^{\Omega}\right)^{\dagger} O_X(\tau) U_{0,t}^{\Omega} \right\| \\
&= \left\| \int_0^t ds \, \frac{d}{ds} \left(U_I^{\dagger}(s) \left(U_{0,t}^{\Omega}\right)^{\dagger} O_X(\tau) U_{0,t}^{\Omega} U_I(s) \right) \right\| \qquad \text{(B.7)} \\
&= \left\| \int_0^t ds \, U_I^{\dagger}(s) \left[\left(U_{0,t}^{\Omega}\right)^{\dagger} O_X(\tau) U_{0,t}^{\Omega}, V_I(s) \right] U_I(s) \right\| \qquad \text{(B.8)} \\
&\leq \int_0^t ds \, \left\| \left[\left(U_{0,t}^{\Omega}\right)^{\dagger} O_X(\tau) U_{0,t}^{\Omega}, V_I(s) \right] \right\| \qquad \text{(B.9)} \\
&\leq \int_0^t ds \, \left\| \left[\left(U_{s,t}^{\Omega}\right)^{\dagger} O_X(\tau) U_{s,t}^{\Omega}, V(s) \right] \right\|. \qquad \text{(B.10)}
\end{aligned}$$

Thus, Lemma 16 follows. □

By substituting $\Lambda \to AB$, $\Omega \to B$, $O_X \to H_{B:C}$, $\tau \to t$, and noting that operators

supported on disjoint subsets commute, we can show using Lemma 16 that

$$\begin{aligned}
\|\delta_{\text{overlap}}\| &= \left\| \left(U_{0,t}^{AB}\right)^\dagger H_{B:C}(t) U_{0,t}^{AB} - \left(U_{0,t}^{B}\right)^\dagger H_{B:C}(t) U_{0,t}^{B} \right\| \\
&\leq \int_0^t ds \left\| \left[\left(U_{s,t}^{B}\right)^\dagger H_{B:C}(t) U_{s,t}^{B}, H_{A:B}(s) \right] \right\| \\
&\leq \sum_{\vec{a}\in A} \sum_{\vec{b},\vec{b}'\in B} \sum_{\vec{c}\in C} \int_0^t ds \left\| \left[\left(U_{s,t}^{B}\right)^\dagger h_{\vec{b},\vec{c}}(t) U_{s,t}^{B}, h_{\vec{a},\vec{b}'}(s) \right] \right\|, \qquad \text{(B.11)}
\end{aligned}$$

where the observables are, in general, evaluated at different times $s \leq t$. We note that while it is necessary to keep track of s, t for completeness, one should pay more attention to the supports of the operators, as they carry useful information about the locality of the system.

In fact, let us pause for a moment to discuss why the right hand side of Eq. (B.11) should be small when ℓ, the distance between A and C, is large. Whenever the supports of $h_{\vec{a},\vec{b}'}(s)$ and $h_{\vec{b},\vec{c}}(t)$ are far from each other, we can bound their commutator norm using a Lieb-Robinson bound for long-range interactions. We use the bound by Gong et al. [29]:

$$\begin{aligned}
&\left\| \left[\left(U_{s,t}^{B}\right)^\dagger h_{\vec{b},\vec{c}}(t) U_{s,t}^{B}, h_{\vec{a},\vec{b}'}(s) \right] \right\| \\
&\leq c e^{v(t-s)} \left\| h_{\vec{b},\vec{c}}(t) \right\| \left\| h_{\vec{a},\vec{b}'}(s) \right\| \left(\frac{1}{(1-\gamma)^\alpha} \frac{1}{r^\alpha} + \frac{1}{e^{\gamma r}} \right), \qquad \text{(B.12)}
\end{aligned}$$

where $r = \text{dist}\left(\{\vec{a}, \vec{b}'\}, \{\vec{b}, \vec{c}\}\right)$ is the distance between the supports, $\gamma \in (0, 1)$ is a constant that can be made arbitrarily close to 1, and c, v are constants that depend only on D.

However, in contrast to short-range interacting systems, here $\vec{b}, \vec{b}'$ run over all possible sites in B, so in principle the distance between the supports of $h_{\vec{a},\vec{b}'}(s)$ and $h_{\vec{b},\vec{c}}(t)$

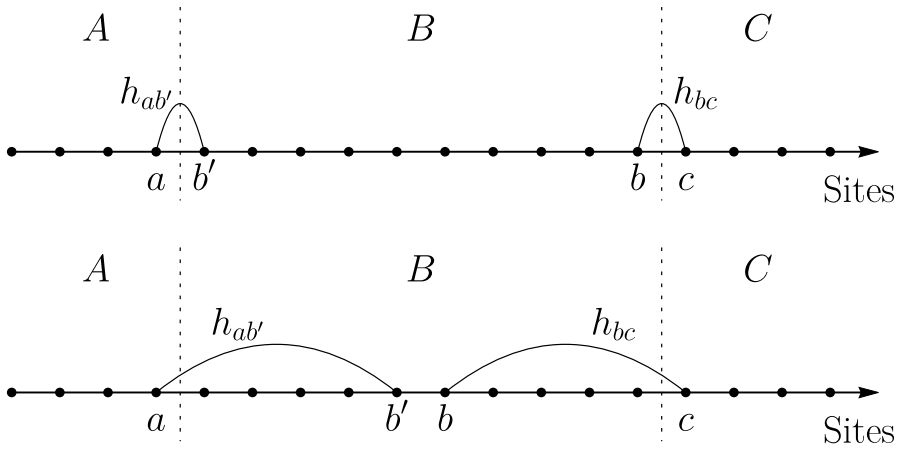

Figure B.1: An illustration of $h_{ab'}$ and h_{bc} in a one-dimensional lattice. For short-range interactions, the sets $\{a, b'\}$ and $\{b, c\}$ are separated by a distance of the same order as the size of B (upper figure). The contributions from these terms to δ_{overlap} are bounded using a Lieb-Robinson bound. However, for long-range interactions, $\{a, b'\}$ and $\{b, c\}$ can be geometrically close to each other (lower figure). In such cases, the norms of $h_{ab'}$ and h_{bc} decay as $|b' - a|^{-\alpha}$ and $|c - b|^{-\alpha}$ and, therefore, their contributions to δ_{overlap} are small.

can be small (Fig. B.1). Fortunately, if that is indeed the case, then although the Lieb-Robinson bound is trivial, the assumption that $\|h_{\vec{a},\vec{b}'}\|$ and $\|h_{\vec{b},\vec{c}}\|$ fall off as $\|\vec{b}' - \vec{a}\|^{-\alpha}$ and $\|\vec{c} - \vec{b}\|^{-\alpha}$, respectively, makes the summand in Eq. (B.11) small.

Let us now evaluate the sum in Eq. (B.11). In the following, we shall consider $\vec{b} \neq \vec{b}'$, since the estimation for the case $\vec{b} = \vec{b}'$ follows a similar, but less complicated argument. Using Gong et al.'s Lieb-Robinson bound for long-range interactions [29]:

$$\begin{aligned}
&\int_0^t ds \| \left[\left(U^B_{s,t}\right)^\dagger h_{\vec{b},\vec{c}}(t) U^B_{s,t}, h_{\vec{a},\vec{b}'}(s) \right] \| \\
&\leq c \int_0^t ds \| h_{\vec{b},\vec{c}}(t) \| \| h_{\vec{a},\vec{b}'}(s) \| \left(\frac{e^{v(t-s)}}{((1-\gamma)r)^\alpha} + \frac{e^{v(t-s)}}{e^{\gamma r}} \right) \\
&\leq \frac{c}{v} \frac{(e^{vt} - 1)}{\|\vec{b} - \vec{c}\|^\alpha \|\vec{a} - \vec{b}'\|^\alpha} \left(\frac{1}{(1-\gamma)^\alpha} \frac{1}{r^\alpha} + \frac{1}{e^{\gamma r}} \right),
\end{aligned} \tag{B.13}$$

where $\gamma \in (0, 1)$ is a constant that can be chosen arbitrarily close to 1, while c, v are finite

and bounded constants for all α, and

$$\begin{aligned} r &= \text{dist}\left(\{\vec{a},\vec{b}'\},\{\vec{b},\vec{c}\}\right) \\ &= \min\left\{\|\vec{b}'-\vec{b}\|,\|\vec{b}'-\vec{c}\|,\|\vec{a}-\vec{b}\|,\|\vec{a}-\vec{c}\|\right\}, \end{aligned} \tag{B.14}$$

is the distance between the supports of $h_{\vec{b},\vec{c}}(t)$ and $h_{\vec{a},\vec{b}'}(s)$ (see Fig. B.1). Since each term of δ_{overlap} contributes a sum of an algebraically decaying as $1/r^{\alpha}$ and an exponential decaying as $e^{-\gamma r}$ terms, it is convenient to evaluate their contributions separately.

First, let us find the contribution from the algebraically decaying part. It is straightforward to find out their contributions to δ_{overlap} when r takes one of the four allowed values. Depending on which value r takes, we use either Lemma 18 or Lemma 20 in Appendix B.7 to evaluate the sums over $\vec{b}$ and $\vec{b}'$. For example, the contribution from the terms where $r = \|\vec{b}'-\vec{b}\|$ is at most

$$\sum_{\vec{a}\in A}\sum_{\vec{b}\neq\vec{b}'\in B}\sum_{\vec{c}\in C}\frac{c(e^{vt}-1)\left(\frac{1}{1-\gamma}\right)^{\alpha}}{v\|\vec{b}-\vec{c}\|^{\alpha}\|\vec{a}-\vec{b}'\|^{\alpha}\|\vec{b}'-\vec{b}\|^{\alpha}}$$

$$\leq \frac{c}{v}\lambda_3\sum_{\vec{a}\in A}\sum_{\vec{b}\in B}\sum_{\vec{c}\in C}\frac{(e^{vt}-1)\left(\frac{2}{1-\gamma}\right)^{\alpha}}{\|\vec{b}-\vec{c}\|^{\alpha}\|\vec{a}-\vec{b}\|^{\alpha}} \tag{B.15}$$

$$\leq \frac{c}{v}\lambda_3\lambda_4\sum_{\vec{a}\in A}\sum_{\vec{c}\in C}\frac{(e^{vt}-1)\left(\frac{4}{1-\gamma}\right)^{\alpha}}{\|\vec{a}-\vec{c}\|^{\alpha}} \tag{B.16}$$

$$\leq \lambda_5(e^{vt}-1)\left(\frac{8}{1-\gamma}\right)^{\alpha}\frac{\Phi(A)}{\ell^{\alpha-D-1}} \tag{B.17}$$

where λ_3, λ_4 are constants that arise after we use Lemma 20 in Appendix B.7 twice to evaluate the sums over $\vec{b}'$ and $\vec{b}$ consecutively, and the sums over $\vec{a}, \vec{c}$ have been bounded

in the previous section (see Eq. (B.2)). The constant λ_5 absorbs both λ_3, λ_4 and the constants from the sums over $\vec{a}, \vec{c}$.

On the other hand, if $r = \|\vec{b}' - \vec{c}\|$, we use Lemma 20 to evaluate the sum over $\vec{b}'$ and Lemma 18 for the sum over $\vec{b}$:

$$\sum_{\vec{a}\in A} \sum_{\vec{b}\neq\vec{b}'\in B} \sum_{\vec{c}\in C} \frac{1}{\|\vec{b}-\vec{c}\|^\alpha} \frac{1}{\|\vec{a}-\vec{b}'\|^\alpha} \frac{c(e^{vt}-1)\left(\frac{1}{1-\gamma}\right)^\alpha}{v\|\vec{b}'-\vec{c}\|^\alpha}$$

$$\leq \frac{c}{v}\lambda_6 \sum_{\vec{a}\in A} \sum_{\vec{b}\in B} \sum_{\vec{c}\in C} \frac{1}{\|\vec{b}-\vec{c}\|^\alpha} \frac{(e^{vt}-1)\left(\frac{2}{1-\gamma}\right)^\alpha}{\|\vec{a}-\vec{c}\|^\alpha} \tag{B.18}$$

$$\leq \frac{c}{v}\lambda_6\lambda_7 \sum_{\vec{a}\in A} \sum_{\vec{c}\in C} \frac{(e^{vt}-1)\left(\frac{2}{1-\gamma}\right)^\alpha}{\|\vec{a}-\vec{c}\|^\alpha} \tag{B.19}$$

$$\leq \lambda_8(e^{vt}-1)\left(\frac{4}{1-\gamma}\right)^\alpha \frac{\Phi(A)}{\ell^{\alpha-D-1}}, \tag{B.20}$$

where λ_6, λ_7 come from the uses of Lemma 20 and Lemma 18 respectively. The constant λ_8 absorbs both λ_6, λ_7 and the constants from the sums over $\vec{a}, \vec{c}$. Repeating for the other values of r, we find that the contribution from the algebraically decaying terms in Equation (B.13) to δ_{overlap} is at most

$$\lambda_9(e^{vt}-1)\left(\frac{8}{1-\gamma}\right)^\alpha \frac{\Phi(A)}{\ell^{\alpha-D-1}}, \tag{B.21}$$

for some constant λ_9.

Next, let us find the contribution from the exponentially decaying term in Equa-

tion (B.13). If $r = \|\vec{b}' - \vec{b}\|$, we have

$$\begin{aligned}
&\sum_{\vec{a}\in A}\sum_{\vec{b}\neq\vec{b}'\in B}\sum_{\vec{c}\in C}\frac{c(e^{vt}-1)}{v\|\vec{b}-\vec{c}\|^{\alpha}\|\vec{a}-\vec{b}'\|^{\alpha}e^{\gamma\|\vec{b}'-\vec{b}\|}} \\
&\leq \frac{c}{v}\lambda_{10}\sum_{\vec{a}\in A}\sum_{\vec{b}\in B}\sum_{\vec{c}\in C}\frac{(e^{vt}-1)}{\|\vec{b}-\vec{c}\|^{\alpha}}\left(\frac{\left(\frac{4}{1-\gamma}\right)^{\alpha}}{\|\vec{a}-\vec{b}\|^{\alpha}}+\frac{\|\vec{a}-\vec{b}\|^{D-1}}{e^{\gamma\|\vec{a}-\vec{b}\|}}\right) \\
&\leq \frac{c}{v}\lambda_{10}\lambda_{11}\sum_{\vec{a}\in A}\sum_{\vec{c}\in C}(e^{vt}-1)\left(\frac{\left(\frac{8}{1-\gamma}\right)^{\alpha}}{\|\vec{a}-\vec{c}\|^{\alpha}}+\frac{\|\vec{a}-\vec{c}\|^{2D-2}}{e^{\gamma\|\vec{a}-\vec{c}\|}}\right) \\
&\leq \lambda_{12}(e^{vt}-1)\Phi(A)\left(\frac{\left(\frac{16}{1-\gamma}\right)^{\alpha}}{\ell^{\alpha-D-1}}+\frac{\ell^{3D-3}}{e^{\gamma\ell}}\right),
\end{aligned} \tag{B.22}$$

where we have applied Lemma 21 in Appendix B.7 to obtain the first inequality, Lemma 21 twice again and Lemma 20 to get the second inequality, and then Lemma 18 and Lemma 19 for the sums over $\vec{a}, \vec{c}$ similarly to Appendix B.1.1. The constants $\lambda_{10}, \lambda_{11}$ arise from the applications of the lemmas and are absorbed into a constant λ_{12}. We note that the constant γ in the last three lines are different from the one in the first line (see Lemma 21 for details). However, they both are constants that can be chosen arbitrarily between 0 and 1. Therefore, we denote them by the same constant γ for convenience.

Repeating the argument for other choices of r in Eq. (B.14), we find that the contribution from the exponentially decaying terms to δ_{overlap} is still at most the right hand side of Eq. (B.22).

Combining Eq. (B.21) and Eq. (B.22), we have

$$\|\delta_{\text{overlap}}\| \leq \lambda_{13}(e^{vt}-1)\Phi(A)\left(\frac{\left(\frac{16}{1-\gamma}\right)^{\alpha}}{\ell^{\alpha-D-1}}+\frac{\ell^{3D-3}}{e^{\gamma\ell}}\right), \tag{B.23}$$

for a constant λ_{13}. Since $\ell^{D-1} \leq \frac{(D-1)!}{\varepsilon^{D-1}} e^{\varepsilon \ell}$ for any arbitrary small positive constant ε, we can upper bound

$$\|\delta_{\text{overlap}}\| \leq c_{\text{ov}}(e^{vt}-1)\Phi(A)\left(\frac{\left(\frac{16}{1-\gamma}\right)^{\alpha}}{\ell^{\alpha-D-1}}+\frac{1}{e^{\gamma\ell}}\right), \tag{B.24}$$

where we have absorbed ε into the definition of γ and c_{ov}. This completes the estimation of δ_{overlap}.

B.2 Error Propagation from Generating Function

In this section, we reproduce a lemma in Ref. [76] which shows how the error in approximating the generating function $\mathcal{G}_W$ propagates to an error of the unitary W_t in Eq. (5.6). Suppose we approximate $\mathcal{G}_W$ by $\mathcal{G}'_W$ such that

$$\|\mathcal{G}_W - \mathcal{G}'_W\| \leq f(t)\delta, \tag{B.25}$$

for some function of time $f(t)$ and δ is time-independent. We shall prove that the unitary W'_t generated by $\mathcal{G}'_W$ approximates W_t with error

$$\|W'_t - W_t\| \leq \delta \int_0^t \mathrm{d}s f(s). \tag{B.26}$$

Proof. By simple differentiation, we have

$$\|W_t^\dagger W_t' - \mathbb{I}\| = \left\| \int_0^t \mathrm{d}s \frac{\mathrm{d}}{\mathrm{d}s}(W_s^\dagger W_s') \right\| \tag{B.27}$$

$$= \left\| \int_0^t \mathrm{d}s W_s^\dagger (G_W - G_W') W_s' \right\| \tag{B.28}$$

$$\leq \int_0^t \mathrm{d}s \|W_s^\dagger (G_W - G_W') W_s'\| \tag{B.29}$$

$$= \int_0^t \mathrm{d}s \|G_W - G_W'\| \tag{B.30}$$

$$\leq \delta \int_0^t \mathrm{d}s f(s). \tag{B.31}$$

□

B.3 Proof of the Lieb-Robinson Bound from Quantum Simulation Algorithms

We present a more detailed proof of Theorem 7 in this section. The key ingredient in the proof of Theorem 7 is the following lemma.

Lemma 17. *Denote by $\mathcal{B}_r = \{\vec{i} \in \Lambda : \|\vec{i}\| \leq r\}$ a D-ball of radius r centered around the origin. Let O_X be an observable supported on $X = \mathcal{B}_{r_0}$ with r_0 being finite. For each $U_{0,T}^\Lambda$ and a positive integer M, there exists a unitary $\tilde{U}$ supported on a D-ball $\mathcal{B}_r$ with $r = r_0 + M\ell$ such that*

$$\left\| \left(U_{0,T}^\Lambda\right)^\dagger O_X U_{0,T}^\Lambda - \tilde{U}^\dagger O_X \tilde{U} \right\| \leq b_1 M e^{vt} \left(r - \ell\right)^{D-1} \xi_\alpha(\ell), \tag{B.32}$$

where b_1 is a constant, $t = T/M$ and $\ell \in (0, R)$ is a free parameter.

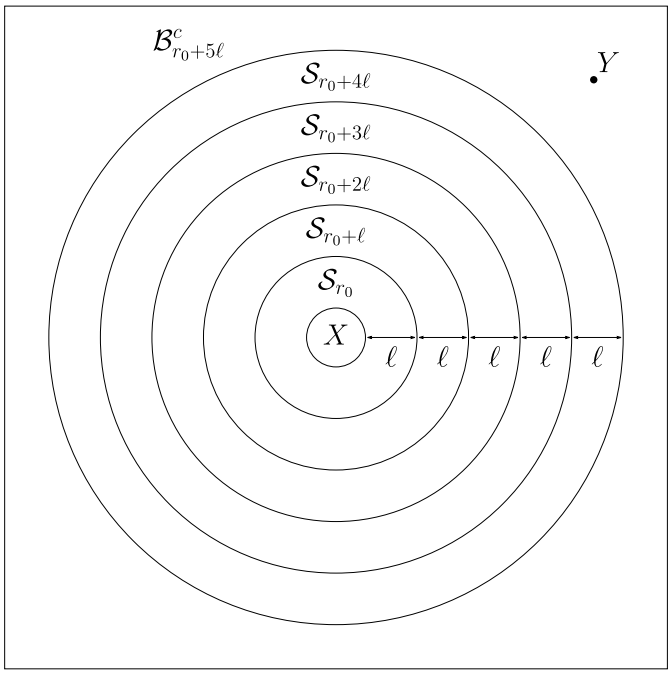

Figure B.2: An example of the subset $X = \mathcal{B}_{r_0}$ and five shells $\mathcal{S}_r$ for $r = r_0, r_0 + \ell, \ldots, r_0 + 4\ell$. The operator O_Y is supported on Y, which lies on $\mathcal{B}^c_{r_0+5\ell}$, the complement of the ball $\mathcal{B}_{r_0+5\ell}$.

Proof. We shall prove the lemma by constructing the unitary $\tilde{U}$. In addition to $\mathcal{B}_r$ above, we define

$$\mathcal{S}_r = \mathcal{B}_{r+\ell} \setminus \mathcal{B}_r \tag{B.33}$$

to be a shell consisting of sites between r and $r + l$ away from the origin (Fig. B.2).

We divide $[0, T]$ into M equal time intervals, namely $[(M-k-1)t, (M-k)t]$ for $k = 0, \ldots, M-1$, where $t = T/M$. The unitary $U^\Lambda_{0,T}$ then naturally decomposes into a product of unitaries $U^\Lambda_k \equiv U^\Lambda_{(M-k-1)t,(M-k)t}$:

$$U^\Lambda_{0,T} = U^\Lambda_0 U^\Lambda_1 \ldots U^\Lambda_{M-1}. \tag{B.34}$$

We now use Lemma 3 to further decompose each U_k^Λ into evolutions of subsystems. We start with $k = 0$ and use Lemma 3 with $A \to X = \mathcal{B}_{r_0}, B \to \mathcal{S}_{r_0}$, and $C \to \mathcal{B}^c_{r_0+\ell}$ (Fig. 5.3) to decompose $\left(U_0^\Lambda\right)^\dagger$ (instead of U_0^Λ):

$$\begin{aligned}&\left\| \left(U_0^\Lambda\right)^\dagger - \left(U_0^{\mathcal{B}_{r_0+\ell}}\right)^\dagger U_0^{\mathcal{S}_{r_0}} \left(U_0^{\mathcal{B}^c_{r_0}}\right)^\dagger \right\| \\ &\leq c_0 e^{vt} \Phi(\mathcal{B}_{r_0}) \xi_\alpha(\ell), \end{aligned} \tag{B.35}$$

where again $\mathcal{B}^c_{r_0}$ denotes the complement subset $\mathbb{R}^D \setminus \mathcal{B}_{r_0}$, and $\Phi(\mathcal{B}_{r_0})$ is the boundary area of $\mathcal{B}_{r_0}$. This choice of decomposition allows us to eliminate the contribution to the evolution from the terms of the Hamiltonian that commute with X, i.e. those supported entirely on $\mathcal{B}^c_{r_0}$. Explicitly, we have:

$$\begin{aligned}&\left(U_0^\Lambda\right)^\dagger O_X U_0^\Lambda \\ &\approx \left(U_0^{\mathcal{B}_{r_0+\ell}}\right)^\dagger U_0^{\mathcal{S}_{r_0}} \left(U_0^{\mathcal{B}^c_{r_0}}\right)^\dagger O_X U_0^{\mathcal{B}^c_{r_0}} \left(U_0^{\mathcal{S}_{r_0}}\right)^\dagger U_0^{\mathcal{B}_{r_0+\ell}} &\text{(B.36)} \\ &= \left(U_0^{\mathcal{B}_{r_0+\ell}}\right)^\dagger O_X U_0^{\mathcal{B}_{r_0+\ell}} = \tilde{U}_0^\dagger O_X \tilde{U}_0, &\text{(B.37)}\end{aligned}$$

where $\tilde{U}_0 \equiv U_0^{\mathcal{B}_{r_0+\ell}}$ is supported entirely on $\mathcal{B}_{r_0+\ell}$.

If we repeat the above argument for U_1^Λ but with O_X replaced by $\tilde{U}_0^\dagger O_X \tilde{U}_0$, we can approximate

$$\left(U_1^\Lambda\right)^\dagger \tilde{U}_0^\dagger O_X \tilde{U}_0 U_1^\Lambda \approx \tilde{U}_1^\dagger \tilde{U}_0^\dagger O_X \tilde{U}_0 \tilde{U}_1, \tag{B.38}$$

for some $\tilde{U}_1$ supported entirely on $\mathcal{B}_{r_0+2\ell}$. The error of this approximation is at most

$c_0 e^{vt} \Phi(\mathcal{B}_{r_0+\ell}) \xi_\alpha(\ell)$.

By induction to all $k = 2, \ldots, M-1$, we can construct $\tilde{U} = \tilde{U}_0 \tilde{U}_1 \ldots \tilde{U}_{T-1}$ such that

$$\left(U_{0,T}^{\Lambda}\right)^\dagger O_X U_{0,T}^{\Lambda} \approx \tilde{U}^\dagger O_X \tilde{U}, \tag{B.39}$$

where the overall error is at most

$$\begin{aligned}
&\sum_{k=0}^{M-1} c_0 e^{vt} \Phi(\mathcal{B}_{r_0+k\ell}) \xi_\alpha(\ell) \\
&\leq M c_0 e^{vt} \Phi(\mathcal{B}_{r_0+(M-1)\ell}) \xi_\alpha(\ell) &\text{(B.40)} \\
&\leq \underbrace{c_0 \frac{2\pi^{\frac{D}{2}}}{\Gamma(\frac{D}{2})}}_{\equiv b_1} M e^{vt} \left(r_0 + (M-1)\,\ell\right)^{D-1} \xi_\alpha(\ell), &\text{(B.41)}
\end{aligned}$$

and where we have replaced the surface area $\Phi(\mathcal{B}_r)$ of a D-ball $\mathcal{B}_r$ by $\frac{2\pi^{\frac{D}{2}}}{\Gamma(\frac{D}{2})} r^{D-1}$ and M by T/t. Also by induction, the unitary $\tilde{U}$ is supported entirely on $\mathcal{B}_{r_0+M\ell}$. Therefore the lemma follows. □

We are now ready to prove our Lieb-Robinson bound in Theorem 7. Without loss of generality, we assume the origin is in X. Since $\|X\| = \mathcal{O}(1)$, there exists $r_0 = \mathcal{O}(1)$ such that X is a subset of $\mathcal{B}_{r_0}$. By Lemma 17, there exists a unitary $\tilde{U}$ supported entirely on a D-ball $\mathcal{B}_r$ with $r = r_0 + M\ell$ such that

$$\begin{aligned}
\varepsilon &= \left\| \left(U_{0,T}^{\Lambda}\right)^\dagger O_X U_{0,T}^{\Lambda} - \tilde{U}^\dagger O_X \tilde{U} \right\| \\
&\leq b_1 M e^{vt} \left(r - \ell\right)^{D-1} \xi_\alpha(\ell). &\text{(B.42)}
\end{aligned}$$

If we choose the number of time slices (M) and the block size (ℓ) such that $r \leq R + r_0$, the set Y will lie outside the support $\mathcal{B}_r$ of $\tilde{U}^\dagger O_X \tilde{U}$, and therefore $\tilde{U}^\dagger O_X \tilde{U}$ will commute with O_Y. Note that for a fixed value of M, the error should decrease with a larger value of ℓ. Therefore, to prove the strongest bound, we should choose ℓ as large as possible, i.e., $\ell = R/M$, and hence $M = R/\ell$. Substituting the value of M and $t = T/M$ into Eq. (B.42), we obtain the error bound in terms of ℓ alone:

$$\varepsilon \leq b_2 \frac{R}{\ell} (e^{\frac{vT\ell}{R}} - 1) \left(1 + R - \ell\right)^{D-1} \xi_\alpha(\ell), \tag{B.43}$$

where $b_2 = b_1 r_0^{D-1}$ is a finite constant. We note that the above bound is valid for all values of $\ell \leq R$. The tightest bound can therefore be obtained by choosing a value for ℓ that minimizes the above expression. Our intuition and numerical evidence suggest that this happens when $\ell \sim \frac{R\alpha}{vT}$, so in the below analysis, we aim to choose ℓ as close to this value as possible.

To proceed, we consider two regimes of time T, when $vT \geq \alpha$ and when $vT < \alpha$. In the former regime, we choose $\ell = \frac{R\alpha}{vT} \leq R$ and substitute into Eq. (B.43) to get

$$\begin{aligned} \varepsilon &\leq b_2 \frac{vT}{\alpha} (e^\alpha - 1) \left(1 + R\left(1 - \frac{\alpha}{vT}\right)\right)^{D-1} \xi_\alpha \left(\frac{R\alpha}{vT}\right) \\ &\leq \underbrace{b_2 v 2^{D-1}}_{\equiv c_{\text{lr}}} T R^{D-1} \xi_\alpha \left(\frac{R\alpha}{vT}\right), \end{aligned} \tag{B.44}$$

where we have used $\frac{e^\alpha - 1}{\alpha} \leq 1$ for all $\alpha \geq 1$, $1 - \frac{\alpha}{vT} \leq 1$ and $1 + R \leq 2R$. In particular, if $\frac{R\alpha}{vT} > x_0$, where x_0 is the larger solution of $x^{\alpha - D - 1} = e^{\gamma x}$, the algebraically decaying

term in ξ_α dominates the exponentially decaying one, and therefore

$$\xi_\alpha\left(\frac{R\alpha}{vT}\right) = \left(\frac{16}{1-\gamma}\right)^\alpha \frac{1}{\left(\frac{R\alpha}{vT}\right)^{\alpha-D-1}} + e^{-\gamma\left(\frac{R\alpha}{vT}\right)}, \tag{B.45}$$

$$\leq 2\left(\frac{16}{1-\gamma}\right)^\alpha \frac{1}{\left(\frac{R\alpha}{vT}\right)^{\alpha-D-1}} \tag{B.46}$$

$$= 2\left(\frac{16}{1-\gamma}\right)^\alpha \left(\frac{v}{\alpha}\right)^{\alpha-D-1}\left(\frac{T}{R}\right)^{\alpha-D-1}. \tag{B.47}$$

Combining Eqs. (B.44) and (B.47), we obtain a bound on the commutator norm:

$$\mathcal{C}(T,R) \leq \varepsilon \leq c_{\mathrm{lr},\alpha}\frac{T^{\alpha-D}}{R^{\alpha-2D}}, \tag{B.48}$$

where

$$c_{\mathrm{lr},\alpha} \coloneqq 2c_{\mathrm{lr}}\left(\frac{16}{1-\gamma}\right)^\alpha\left(\frac{v}{\alpha}\right)^{\alpha-D-1}. \tag{B.49}$$

The light cone implied by the bound is

$$T \gtrsim R^{\frac{\alpha-2D}{\alpha-D}}. \tag{B.50}$$

In the limit $\alpha \to \infty$, the exponent of the light cone converges to one at a rate given by

$$\mu = \lim_{\alpha\to\infty}\frac{\left|\frac{\alpha+1-2D}{\alpha+1-D}-1\right|}{\left|\frac{\alpha-2D}{\alpha-D}-1\right|} = 1. \tag{B.51}$$

On the other hand, if $vT < \alpha$, we simply choose $\ell = R$. Equation (B.43) then

becomes

$$\mathcal{C}(T,R) \leq \varepsilon \leq b_2(e^{vT}-1)\xi_\alpha(R). \tag{B.52}$$

Therefore, we arrive at the Lieb-Robinson bound in Theorem 7 with $\tilde{c}_{\text{lr}} = b_2$.

B.4 Proof of Lemma 4 in Higher Dimensions

In this section, we discuss the construction of the circuit in Lemma 4 that generalizes the lemma to higher dimensions. Similar to the $D = 1$ case, we first break the unitary into $\mathcal{O}(T)$ unitaries $\exp(-iHt)$ for some $t = \mathcal{O}(1)$. We then use an algorithm consisting of D steps to break the simulation of $\exp(-iHt)$ into simulations of Hamiltonians on smaller hypercubes of size at most 2ℓ. In the first of the D steps, we cut the D-dimensional lattice into L/ℓ layers, each with the same thickness ℓ, a parameter to be chosen later. In this step, the cross section of the cut is L^{D-1}. Therefore, by Lemma 3, each time a new layer is generated, we accumulate an error of

$$\mathcal{O}\left(\frac{L^{D-1}}{\ell^{\alpha-D-1}}\right).$$

For TL/ℓ layers of the first step, the accumulated error will be $\varepsilon^{(1)} = \mathcal{O}\left(TL^D/\ell^{\alpha-D}\right)$.

Next, for each of the $\mathcal{O}(TL/\ell)$ layers of $D-1$ dimensions, we break them again into L/ℓ layers of $D-2$ dimensions. Using Lemma 3 with a cross section L^{D-2}, we find

the error of the second step

$$\varepsilon^{(2)} = \frac{TL}{\ell}\frac{L}{\ell}\mathcal{O}\left(\frac{L^{D-2}}{\ell^{\alpha-(D-1)-1}}\right) = \mathcal{O}\left(\frac{TL^D}{\ell^{\alpha-D+2}}\right), \tag{B.53}$$

which decreases with ℓ faster than the error of the first step.

More explicitly, in the kth of the D steps, the error is $\varepsilon^{(k)} = \mathcal{O}\left(L^D/\ell^{\alpha-D-2k}\right)$, which is dominated by the error in the first step for all $k > 1$. Therefore, the error of cutting the D-dimensional lattice of size L into L^D/ℓ^D subsystems is still $\mathcal{O}\left(TL^D/\ell^{\alpha-D}\right)$. To meet a fixed total error ε, we need to choose $\ell \propto \left(TL^D/\varepsilon\right)^{\frac{1}{\alpha-D}}$. The geometrical constraint $\ell < L$ requires $\alpha > 2D$. Finally, simulating each of the $\mathcal{O}\left(TL^D/\ell^D\right)$ subsystems using the LCU algorithm up to $\varepsilon\ell^D/(TL^D)$ accuracy requires $\mathcal{O}\left(\ell^{3D}\log\left(TL^D/\varepsilon\ell^D\right)\right)$ quantum gates. Therefore, the overall gate complexity of the algorithm is

$$G_D = \mathcal{O}\left(\frac{(Tn)^{1+\frac{2D}{\alpha-D}}}{\varepsilon^{\frac{2D}{\alpha-D}}}\log\frac{Tn}{\varepsilon}\right). \tag{B.54}$$

B.5 Estimation of the Actual Gate Count

In this section, we describe how we estimate the actual gate count of the HHKL algorithm and the QSP algorithm in simulating one-dimensional power-law systems.

The direct implementation of the QSP algorithm requires computing a sequence of rotation angles on a classical computer, which is prohibitive for large-size Hamiltonian simulation. Instead, we use a suboptimal approach described in Ref. [7]. To simulate $H = \sum_{j=1}^{L}\beta_j H_j$ for time t and accuracy ε, where L is the number of terms in the Hamiltonian,

$\beta_j \geq 0$ and H_j are both unitary and Hermitian, we divide the entire evolution into r segments. We choose r sufficiently large so that each segment is short enough for the classical preprocessing. Specifically, we choose

$$r = \left\lceil \frac{\sum_j \beta_j t}{\tau_{\max}} \right\rceil \tag{B.55}$$

and $\tau_{\max} = 1000$ [76]. Within each segment, we choose q to be the smallest positive integer satisfying

$$\frac{4(\sum_j \beta_j t/r)^q}{2^q q!} \leq \frac{\varepsilon}{8r}, \tag{B.56}$$

so that the overall error is at most ε. This gives $M = 2(q-1)$ phased iterates within each segment [7].

The number of elementary operations of each phased iterate is $\log(L) + 4L + 8L$. Here, the first term corresponds to the reflection along an L-dimensional state $|0\rangle$; the second term costs the preparation/unpreparation of an L-dimensional state; and the third term is the cost of selecting L two-body operators. We thus estimate the gate complexity of the QSP algorithm as $\big(\log(L) + 12L\big) rM$.

Next, in order to determine the gate count of the HHKL algorithm, we need an estimate for the error of the unitary decomposition in Lemma 3. Recall that for $D = 1$, the error given by our analysis is $b/\ell^{\alpha-2}$, where b is a constant that can be estimated numerically by computing the actual error for small values of ℓ and extrapolating for larger ℓ.

Since simulating the evolution of a generic system is classically intractable even for

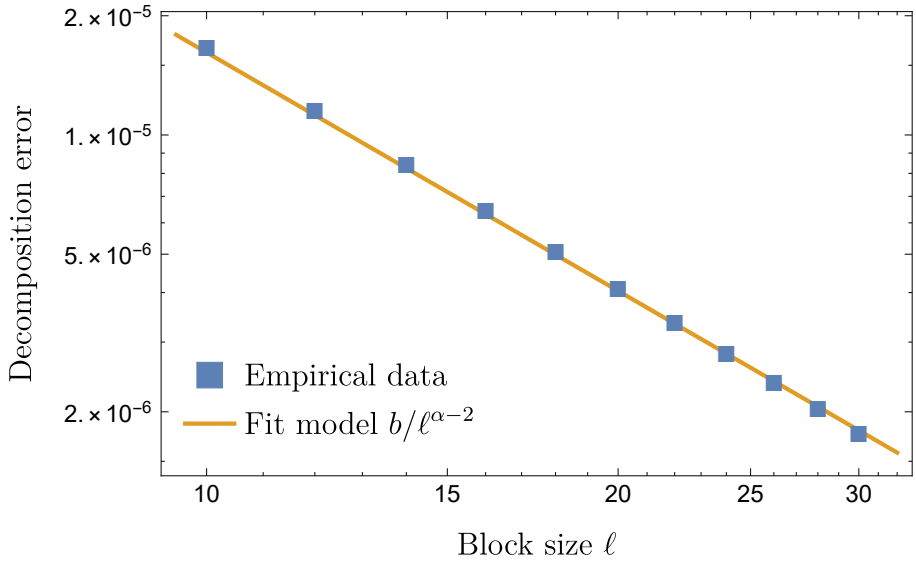

Figure B.3: The empirical error of the unitary decomposition in Lemma 3, computed for the single-excitation one-dimensional Heisenberg chain ($\alpha = 4$) in Eq. (7.18) at different values of ℓ. The system size is fixed at $n = 300$ and the evolution time at $t = 0.01$. We fit the data (blue square) to the theoretical model $b/\ell^{\alpha-2}$ and obtain $b = 1.62 \times 10^{-3}$.

a moderate system size, we study only the one-dimensional Heisenberg model given in Eq. (7.18) and restrict our calculation to the single-excitation subspace. In Fig. B.3, we plot the error of the unitary decomposition in Lemma 3 at several different values of ℓ (for system size $n = 300$ and evolution time $t = 0.01$). The scaling of the error agrees well with our prediction. By fitting the data to $b/\ell^{\alpha-2}$, we obtain an estimate $b = 1.62 \times 10^{-3}$.

Recall that there are T/t time slices in the HHKL algorithm. In each time slice, there are n/ℓ blocks of size ℓ and $2n/(2\ell)$ blocks of size 2ℓ. To meet the total error at most ε, we need to choose (see also Eq. (5.20))

$$\ell = \left(\frac{T}{t} \frac{2nb}{\varepsilon} \right)^{\frac{1}{\alpha-1}}. \tag{B.57}$$

By multiplying the number of blocks by the gate count for using QSP to simulate a single

block, we arrive at the total gate count presented in Fig. 5.5.

B.6 Numerical Performance of the Product Formula

This section includes the numerical performance of the fourth-order product formula (PF4) used to simulate the evolution of the system given in Eq. (7.18) for time $T = n$. We plot this numerical performance as well as the theoretical estimates for the gate counts of the PF4, LCU, QSP, and HHKL algorithms in Fig. B.4.

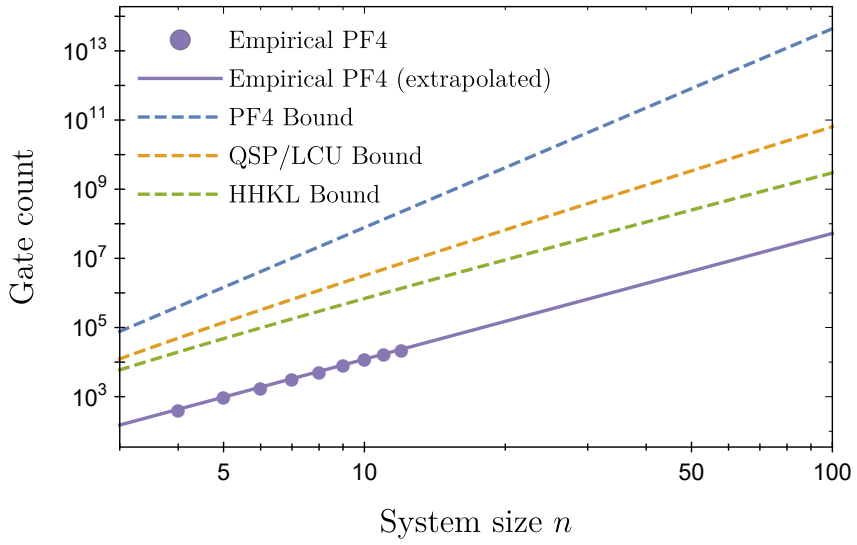

Figure B.4: The empirical gate count of PF4 (purple dots) from $n = 4$ to $n = 12$, extrapolated to larger system sizes (solid, purple), for simulating the dynamics of the Hamiltonian in Eq. (7.18) for time $T = n$ at a fixed error tolerance. The error bars are smaller than the size of the markers and hence not visible in the plot. Also shown in dashed lines are the *slopes* of the gate counts of several advanced algorithms for comparison. These slopes represent the scaling of the gate counts as functions of n. Their y-intercepts, which represent a constant multiplicative factor, should be ignored.

B.7 Mathematical Tools

This section contains a collection of mathematical results omitted from the previous sections. In Appendix B.7.1, we present the upper bounds on standard sums we use in the proof of Lemma 3 in Appendix B.1. In Appendix B.7.2, we show how we estimate the sum over the convex set A in Eq. (B.3) by parameterizing the elements of the set by their distance to the boundary of A. We also note that we use the same notation "λ" for constants that appear in different lemmas.

B.7.1 Standard Sums

In this section, we present upper bounds on a a few standard sums used in the previous sections. Specifically, we use Lemma 18 to bound Eq. (B.2), Eq. (B.18), Lemma 19 to bound Eq. (B.15), Eq. (B.22), Lemma 20 to bound Eq. (B.18), Eq. (B.22), and Lemma 21 to bound Eq. (B.22).

Lemma 18. *Let Λ be a D-dimensional lattice and $\vec{r}$ be the coordinates of sites in Λ. For $\alpha > D+1$ and $R > \sqrt{D}$, there exists a constant λ that may depend on D but not on R, α such that:*

$$\sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}} \frac{1}{\|\vec{r}\|^{\alpha}} \leq \frac{\lambda}{(R-\sqrt{D})^{\alpha-D}}. \tag{B.58}$$

In particular, it implies that the sum $\sum_{\vec{r}\in\Lambda}$ converges for all $\alpha > D$.

Proof. The proof of this bound is straightforward. For simplicity, we first assume none

of the coordinates of $\vec{r}$ is zero. Since $\frac{1}{x^\alpha}$ is a decreasing function of x for all $\alpha > 0$, we can always bound the sum over such $\vec{r}$ by an integral

$$\begin{aligned}\sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}}{}' \frac{1}{\|\vec{r}\|^\alpha} &\leq \int_{\|\vec{r}\|\geq R-\sqrt{D}} \frac{\mathrm{d}^D\vec{r}}{\|\vec{r}\|^\alpha} \\ &= \frac{2\pi^{\frac{D}{2}}}{\Gamma(\frac{D}{2})}\int_{R-\sqrt{D}}^{\infty}\frac{\mathrm{d}r}{r^{\alpha-D+1}} \leq \frac{g(D)}{(R-\sqrt{D})^{\alpha-D}},\end{aligned} \tag{B.59}$$

where $\sum'$ denotes the sum over $\vec{r}$ with no zero coordinate and $g(D) \equiv 2\pi^{\frac{D}{2}}\big/\Gamma(\frac{D}{2})$.

Next, consider $\vec{r}$ with exactly one zero coordinate. These sites lie on D hyperplanes, each of dimension $(D-1)$. Therefore the contribution from them can be evaluated using the above integral with $D \to D-1$:

$$\frac{Dg(D-1)}{(R-\sqrt{D-1})^{\alpha-D+1}} < \frac{Dg(D-1)}{(R-\sqrt{D})^{\alpha-D}}. \tag{B.60}$$

By repeating this argument for the sums over $\vec{r}$ with different number of zero coordinates, we arrive at

$$\sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}} \frac{1}{\|\vec{r}\|^\alpha} \leq \frac{\lambda}{(R-\sqrt{D})^{\alpha-D}}, \tag{B.61}$$

where $\lambda = \sum_{d=0}^{D}\binom{D}{d}g(D-d)$ is a constant independent of R. □

Lemma 19. *Let Λ be a D-dimensional lattice and $\vec{r}$ be the coordinates of sites in Λ. For*

all $R > 0$, there exists a constant λ that may depend on β, D but not on R such that:

$$\sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}} \frac{\|\vec{r}\|^{\beta}}{e^{\|\vec{r}\|}} \leq \frac{\lambda R^{\beta+D-1}}{e^{R}}, \tag{B.62}$$

where β is a positive constant. In particular, it also implies that the sum $\sum_{\vec{r}\in\Lambda}$ converges.

Proof. The proof of this lemma follows the same idea as of Lemma 18. However, note that the function $x^{\beta}e^{-x}$ is a decreasing function of x only when $x \geq x_0$ for some x_0 that depends only on β. Therefore, if $R \geq x_0$, we follow the exact same lines as in the proof of Lemma 18. For example, if none of the coordinates of $\vec{r}$ is zero, we can bound

$$\begin{aligned}
\sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}}{}' \frac{\|\vec{r}\|^{\beta}}{e^{\|\vec{r}\|}} &\leq \int_{\|\vec{r}\|\geq R-\sqrt{D}} \frac{\|\vec{r}\|^{\beta}}{e^{\|\vec{r}\|}} \mathrm{d}^{D}\vec{r} \\
&= \frac{2\pi^{\frac{D}{2}}}{\Gamma(\frac{D}{2})} \int_{R-\sqrt{D}}^{\infty} \frac{r^{\beta+D-1}\mathrm{d}r}{e^{r}} \qquad \text{(B.63)} \\
&\leq \lambda_1 \frac{(R-\sqrt{D})^{\beta+D-1}}{e^{R-\sqrt{D}}} \leq \lambda_2 \frac{R^{\beta+D-1}}{e^{R}}, \qquad \text{(B.64)}
\end{aligned}$$

for some constants λ_1, λ_2 that depend only on β, D.

On the other hand, if $R < x_0$, we consider

$$\lambda = \max\left\{ \frac{e^{R}}{R^{\beta+D-1}}, \sum_{\substack{\vec{r}\in\Lambda \\ \|\vec{r}\|\geq R}} \frac{\|\vec{r}\|^{\beta}}{e^{\|\vec{r}\|}} \right\}. \tag{B.65}$$

The lemma should follow if we can argue that λ can be chosen independently of R. Indeed, since $1 \leq R < x_0$ and from the previous calculation, we know that the sum over $\vec{r}$ converges to a constant that depends only on β, D. This concludes the proof of

Lemma 19. □

Lemma 20. *Let $\vec{a}, \vec{b}, \vec{c}$ be three distinct sites in a D-dimensional lattice Λ. For all $\alpha > D$,*

$$\sum_{\vec{b}\in\Lambda} \frac{1}{\|\vec{a}-\vec{b}\|^\alpha} \frac{1}{\|\vec{b}-\vec{c}\|^\alpha} \leq \frac{\lambda 2^\alpha}{\|\vec{a}-\vec{c}\|^\alpha}, \tag{B.66}$$

where λ is a constant independent of $\vec{a}, \vec{c}, \alpha$.

Proof. A proof of the lemma is presented in Ref. [18]. □

Lemma 21. *Let $\vec{a}, \vec{b}, \vec{c}$ be three distinct sites in a D-dimensional lattice Λ. For all $\alpha > D$, $\gamma \in (0,1)$, and positive integers $\beta \in \mathbb{Z}^+$, there exists a constant $\gamma' \in (0,1)$ such that*

$$\sum_{\vec{b}\in\Lambda} \frac{1}{\|\vec{a}-\vec{b}\|^\alpha} \frac{\|\vec{b}-\vec{c}\|^\beta}{e^{\gamma\|\vec{b}-\vec{c}\|}} \leq \frac{\lambda\left(\frac{4}{1-\gamma'}\right)^\alpha}{\|\vec{a}-\vec{c}\|^\alpha} + \frac{\lambda'\|\vec{a}-\vec{c}\|^{\beta+D-1}}{e^{\gamma'\|\vec{a}-\vec{c}\|}}, \tag{B.67}$$

where λ, λ' are constants that may depend on β, D, but not on $\vec{a}, \vec{c}, \alpha$.

Proof. Without loss of generality, assume $\vec{c} = 0$. Let $\ell = \|\vec{c}-\vec{a}\| = \|\vec{a}\|$ be the distance between $\vec{c}$ and $\vec{a}$. We need to prove

$$\sum_{\vec{b}\in\Lambda} \frac{1}{\|\vec{a}-\vec{b}\|^\alpha} \frac{\|\vec{b}\|^\beta}{e^{\gamma\|\vec{b}\|}} \leq \frac{\lambda\left(\frac{4}{1-\gamma'}\right)^\alpha}{\ell^\alpha} + \frac{\lambda'\ell^{\beta+D-1}}{e^{\gamma'\ell}}. \tag{B.68}$$

Let $\mathcal{B}_{\mu\ell}$ be a D-ball of radius $\mu\ell$ centered around $\vec{c}$ for some arbitrary constant $\mu \in (0,1)$. We shall divide the sum over $\vec{b}$ into two regimes, corresponding to $\vec{b}$ inside and outside $\mathcal{B}_{\mu\ell}$.

In the first regime where $\vec{b}$ is inside $\mathcal{B}_{\mu\ell}$, we can show using the triangle inequality

that $\|\vec{a}-\vec{b}\| \geq (1-\mu)\ell$. Therefore, the sum over these $\vec{b}$ can be bounded by

$$\frac{1}{((1-\mu)\ell)^{\alpha}} \sum_{\vec{b}\in\mathcal{B}_{\mu\ell}} \frac{\|\vec{b}\|^{\beta}}{e^{\|\vec{b}\|}} \leq \frac{\lambda\left(\frac{2}{1-\mu}\right)^{\alpha}}{\ell^{\alpha}}, \tag{B.69}$$

where we have used the fact that $\sum_{\vec{b}\in\mathcal{B}_{\mu\ell}} \frac{\|\vec{b}\|^{\beta}}{e^{\|\vec{b}\|}}$ converges and is bounded by a constant λ which may depend only on D, β.

In the second regime, we bound $\|\vec{a}-\vec{b}\| \geq 1$ to obtain

$$\begin{aligned}\sum_{\vec{b}\notin\mathcal{B}_{\mu\ell}} \frac{1}{\|\vec{a}-\vec{b}\|^{\alpha}} \frac{\|\vec{b}\|^{\beta}}{e^{\gamma\|\vec{b}\|}} &\leq \sum_{\vec{b}\notin\mathcal{B}_{\mu\ell}} \frac{\|\vec{b}\|^{\beta}}{e^{\gamma\|\vec{b}\|}} \\ &\leq \lambda' \frac{\ell^{\beta+D-1}}{2^{\alpha} e^{\gamma\mu\ell}},\end{aligned} \tag{B.70}$$

where the last sum is bounded using Lemma 19 and noting that $\mu < 1$.

Combining Eq. (B.69), Eq. (B.70), we arrive at a bound

$$\sum_{\vec{b}\in\Lambda} \frac{1}{\|\vec{a}-\vec{b}\|^{\alpha}} \frac{\|\vec{b}\|^{\beta}}{e^{\gamma\|\vec{b}\|}} \leq \frac{\lambda\left(\frac{2}{1-\mu}\right)^{\alpha}}{\ell^{\alpha}} + \frac{\lambda'\ell^{\beta+D-1}}{e^{\gamma\mu\ell}}. \tag{B.71}$$

Let $\gamma' = \gamma\mu$ and take $\mu \leq \frac{1}{2-\gamma}$, it is straightforward to show that $\frac{2}{1-\mu} \leq \frac{4}{1-\gamma'}$, and therefore,

$$\sum_{\vec{b}\in\Lambda} \frac{1}{\|\vec{a}-\vec{b}\|^{\alpha}} \frac{\|\vec{b}\|^{\beta}}{e^{\gamma\|\vec{b}\|}} \leq \frac{\lambda\left(\frac{4}{1-\gamma'}\right)^{\alpha}}{\ell^{\alpha}} + \frac{\lambda'\ell^{\beta+D-1}}{e^{\gamma'\ell}}. \tag{B.72}$$

Note that if we choose $\mu = \frac{1}{2-\gamma}$, then $\gamma' = \frac{\gamma}{2-\gamma}$ takes on a value between 0 and 1, which can be arbitrarily close to 1. □

B.7.2 Parameterizing a Convex Set

In this subsection, we show how we evaluate the sum over $\vec{a}$ in Eq. (B.3). First, we parameterize a convex set by the distance to its boundary. The following lemma simplifies a sum over every site in a convex set to a sum over the above distance, multiplied by the boundary area of the set.

Lemma 22. *Let $A \subset \mathbb{R}^D$ be a compact and convex set in $\mathbb{R}^D$ with non-empty interior. Let $C \subset \mathbb{R}^D$ be another subset disjoint from A, and let $\ell = \mathrm{dist}\,(A, C)$ be the smallest distance between elements of the two sets. Furthermore, we denote by $\ell_{\vec{a}} = \mathrm{dist}\,(\vec{a}, C)$ the minimal distance from a given lattice site $\vec{a}$ in A to C. For a decreasing function $f : \mathbb{R} \to \mathbb{R}$, we shall have*

$$\sum_{\vec{a} \in A \cap \Lambda} f(\ell_{\vec{a}}) \leq 2\eta \Phi(A) \sum_{\mu=0}^{\infty} f(\ell + \mu), \tag{B.73}$$

where η is a constant that may depend only on D and $\Phi(A)$ is the boundary area of A.

Proof. Let us divide the set $A \in \mathbb{R}^D$ into disjoint subsets

$$S_\mu = \{\vec{a} \in A : \mu \leq \mathrm{dist}\,(\vec{a}, \partial A) \leq \mu + 1\} \tag{B.74}$$

for $\mu = 0, 1, \ldots$ Note that the assumption that the interior of A is non-empty implies that $\mathrm{dist}\,(\vec{a}, \partial A)$ is not uniformly zero. Roughly speaking, S_μ contains the sites in A whose distances to the boundary ∂A are between μ and $\mu + 1$. Therefore, $\ell_{\vec{a}} \geq \ell + \mu$ for all

$\vec{a} \in S_\mu$. We then obtain

$$\sum_{\vec{a}\in A\cap\Lambda} f(\ell_{\vec{a}}) = \sum_{\mu=0}^{\infty} \sum_{\vec{a}\in S_\mu\cap\Lambda} f(\ell_{\vec{a}}) \tag{B.75}$$

$$\leq \sum_{\mu=0}^{\infty} f(\ell+\mu)\,|S_\mu \cap \Lambda|\,, \tag{B.76}$$

where $|S_\mu \cap \Lambda|$ is the number of lattice sites that lie within S_μ.

Let $A_\mu = \{\vec{a} \in A : \mathrm{dist}\,(\vec{a}, \partial A) \geq \mu\}$ be a subset of A containing sites at least a distance μ from the boundary of A. Clearly, $S_\mu = (A_\mu \setminus A_{\mu+1}) \cup \partial A_{\mu+1}$ and $\partial S_\mu = \partial A_\mu \cup \partial A_{\mu+1}$. Roughly speaking, S_μ is a shell with the outer surface A_μ, the inner surface $A_{\mu+1}$ and a unity thickness. The number of lattice sites in S_μ will be bounded by $\eta\Phi(S_\mu) = \eta(\Phi(A_\mu) + \Phi(A_{\mu+1}))$ (see Appendix B.7.2.1 for the definition of the constant η). Since A is compact and convex, $\Phi(A_{\mu+1}) < \Phi(A_\mu) < \Phi(A)$ (see Appendix B.7.2.2). Therefore, we arrive at the lemma. □

B.7.2.1 The number of lattice sites in a compact region

In this subsection, we shall provide an upper bound on the number of lattice sites inside a compact set $A \subset \mathbb{R}^D$. We use this bound in Eq. (B.76) to estimate the number of lattice sites in the set $|S_\mu \cap \Lambda|$ by its boundary area. Let $A_> = \left\{\vec{a} \in A \cap \Lambda : \mathrm{dist}\,(a, \partial A) > \frac{1}{3}\right\}$ be the set of lattice sites that are at least a distance $\frac{1}{3}$ away from the boundary ∂A, and let $A_\leq = A \setminus A_>$ be the other lattice sites of A.

First, note that for every lattice site $\vec{a}$ in $A_>$, there exists a D-ball $\mathcal{B}_{1/4}(\vec{a})$ of radius $\frac{1}{4}$ that contains no other lattice site and $\mathcal{B}_{1/4}(\vec{a}) \subset A$. Therefore, the number of lattice

sites in $A_>$ is at most $\mathcal{V}(A)/\mathcal{V}(\mathcal{B}_{1/4}(\vec{a})) = \eta_1 \mathcal{V}(A)$, where $\mathcal{V}(A)$ is the volume of A in $\mathbb{R}^D$ and $\eta_1 = \mathcal{V}(\mathcal{B}_{1/4}(\vec{a}))^{-1}$.

Next, to count the lattice sites in $A_\leq$, we note that for every $\vec{a} \in A_\leq$, we can select a point $f(\vec{a}) \in \partial A$ on the boundary such that $\|f(\vec{a}) - \vec{a}\| \leq \frac{1}{3}$. We now argue that $\|f(\vec{a}) - f(\vec{b})\| \geq \frac{1}{3}$ for all distinct lattice sites $\vec{a} \neq \vec{b}$ in $A_\leq$. Indeed, since $\vec{a}, \vec{b}$ are distinct lattice sites, the least distance between them is 1, i.e. $\|\vec{a} - \vec{b}\| \geq 1$. Using a triangle inequality, we can show that

$$\begin{aligned} \|f(\vec{a}) - f(\vec{b})\| &\geq \|\vec{a} - \vec{b}\| - \|f(\vec{a}) - \vec{a}\| - \|f(\vec{b}) - \vec{b}\| \\ &\geq 1 - \frac{1}{3} - \frac{1}{3} = \frac{1}{3}. \end{aligned} \tag{B.77}$$

Therefore, a D-ball $\mathcal{B}_{1/6}(f(\vec{a}))$ around $f(\vec{a}) \in \partial A$ shall contain no $f(\vec{b})$ of any other lattice site $\vec{b} \in A_\leq$. Therefore, the number of lattice sites in $A_\leq$ is at most $\eta_2 \Phi(A)$, where $\Phi(A) = |\partial A|$ is the boundary area of A and η_2 is the area of a $(D-1)$-dimensional disk of radius $1/6$.

In summary, the number of lattice sites in A is therefore at most $\eta_1 V(A) + \eta_2 \Phi(A)$. In particular, for a shell A whose volume $\mathcal{V}(A)$ can be upper bounded by $\eta_3 \Phi(A)$, the number of lattice sites will be at most $\eta \Phi(A)$, where $\eta = \eta_1 \eta_3 + e\eta_2$.

B.7.2.2 Convex sets in $\mathbb{R}^D$ are shrinkable

In the proof of Lemma 22 [see the discussion after Eq. (B.76)], we used the fact that $\Phi(A_\mu) < \Phi(A)$. In this section, we will show that this property of A—which we term *shrinkability*—holds if A belongs to the class of convex and compact sets in $\mathbb{R}^D$.

The formal definition is as follows:

Definition 2 (Shrinkable set)**.** *A compact set $A \subset \mathbb{R}^D$ with boundary ∂A is* shrinkable *if, for all $r > 0$, $A_r = \{\vec{a} \in A : dist\,(\vec{a}, \partial A) \geq r\}$, we have that $\Phi(A_r) = |\partial A_r| \leq |\partial A| = \Phi(A)$.*

In other words, a set is shrinkable if the surface area of the boundary of $A_r \subseteq A$ is no larger than that A. In this section, we will prove that convexity is a sufficient condition for shrinkability. Recall that a set is compact if it is both closed and bounded, whereas convexity is usually defined as follows:

Definition 3. *A set A is convex if for any $x, y \in A$ and any θ such that $0 \leq \theta \leq 1$, we have $\theta x + (1 - \theta)y \in A$.*

Examples of convex sets include D-balls and hyperrectangles, which are also shrinkable. To prove this holds in general, we will first show that if A is convex, then A_r is also convex (or empty) for all $r > 0$. To do this, we formulate an equivalent definition of a convex set as an intersection of halfspaces.

Definition 4. *A halfspace $\mathcal{H}$ is given by the points $\{x \in \mathbb{R}^D \mid a^T x \geq b\}$, where $a \in \mathbb{R}^D \backslash \{0\}$.*

From this definition, it follows that halfspaces are convex sets. A folk lemma [201] states that a closed set A is convex iff

$$A = \bigcap_{k \in I} \{\mathcal{H}_k \mid \mathcal{H}_k \text{ halfspace}, A \subseteq \mathcal{H}_k\},$$

for some countable index set I. In other words, A is equivalent to the intersection of all halfspaces that contain it. Since convexity is preserved under arbitrary intersection, this implies that A is convex. The converse follows from the separating hyperplane theorem—see [201] for details.

With this equivalent definition of convexity in hand, we will prove that A_r is also convex.

Lemma 23. *If a compact set $A \subset \mathbb{R}^D$ is convex, then $A_r = \{\vec{a} \in A : dist\,(\vec{a}, \partial A) \geq r\}$ is convex (or empty) for all $r > 0$.*

Proof. Write A as the intersection of half-spaces $\mathcal{H}_k = \{x \in \mathbb{R}^D \mid a_k^T x \geq b_k\}$, for $k \in I$. Then A_r is the intersection of the half-spaces given by $H_k^r = \{x \in \mathbb{R}^D \mid a_k^T x \geq b_k + r\}$. By the converse of the above lemma, A_r is convex (or empty). □

To show that A is shrinkable, we must show that $\Phi(A_r) = |\partial A_r| \leq |\partial A| = \Phi(A)$. Following a standard technique in the literature, we define the *nearest-point projection* of $\mathbb{R}^D$ onto a convex set and then show that it is a contraction. The following lemma implies that such a mapping is well-defined.

Lemma 24. *Given a non-empty, compact and convex set $A \subseteq \mathbb{R}^D$ and a point $x \in \mathbb{R}^D$, there exists a unique point $p_A(x) \in A$ such that*

$$p_A(x) = \arg\min_{y \in A} \|x - y\|.$$

Proof. Since A is compact, the continuous function $d_x(y) = \|x - y\|$ must achieve its minimum value on A.

Now suppose that minimum value of d_x occurs at a point $y \in A$. We will show that y is unique. Assume for the sake of contradiction that there exists some point $\tilde{y} \in A$ such that $d_x(y) = d_x(\tilde{y})$, but $y \neq \tilde{y}$. Then the set of points x, y, and $\tilde{y}$ form an isosceles triangle, with $\overline{y\tilde{y}}$ as the base. Dropping an altitude from x intersects this line segment at the midpoint m such that $\|x - m\| < \|x - y\| = \|x - \tilde{y}\|$. But m is a convex combination of y and $\tilde{y}$, i.e. $m = \frac{1}{2}(y + \tilde{y}) \in A$, so we have reached a contradiction. Thus, y must be unique, and, therefore, $p_A(x)$ is well-defined. □

The projection function $p_A(x)$ can be interpreted as generalizing the concept of the orthogonal projection into an affine subspace. It is also well-known that the nearest point projection p_A is a contraction mapping.

Lemma 25. *Given a nearest-point projection $p_A : \mathbb{R}^D \to A$ onto a convex set A, it holds for all $x, y \in \mathbb{R}^D$ that*

$$\|p_A(x) - p_A(y)\| \leq \|x - y\|.$$

Proof. While the lemma can be proved for all $x, y \in \mathbb{R}^D$, for our purposes, we only need to consider $x, y \notin A$. Assume that $p_A(x) \neq p_A(y)$. Then consider the hyperplanes H_x and H_y that pass through $p_A(x)$ and $p_A(y)$ respectively, and are perpendicular to the line segment $\overline{p_A(x)p_A(y)}$. (See the geometric diagram in Fig. B.5.)

We prove by contradiction that x (y) and $p_A(y)$ ($p_A(x)$) lie on opposite sides of H_x (H_y). Suppose without loss of generality that x and $p_A(y)$ lie on the same side of H_x. Then the point where the altitude from x intersects the line segment $\overline{p_A(x)p_A(y)}$ would

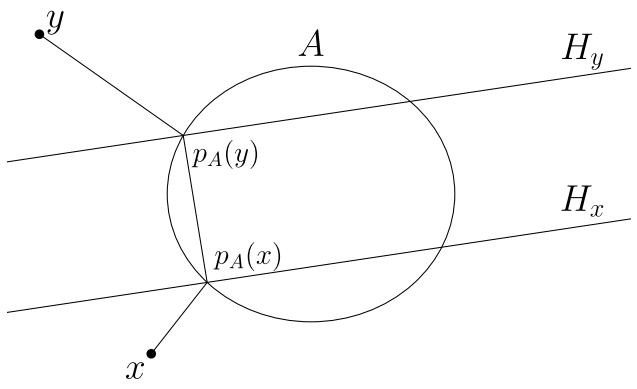

Figure B.5: The nearest-point projection p_A of two points x and y onto a compact set A (oval). Also depicted are the line segment connecting the two image points $p_A(x)$ and $p_A(y)$, as well as the two hyperplanes orthogonal to it.

lie in A, contradicting the fact that $p_A(x)$ is the nearest-point in A to x. Thus, x (y) must lie on the opposite side of H_x (H_y) from $p_A(y)$ ($p_A(x)$). Then, as shown in Fig. B.5, the points x and y must fall outside the rectangular strip between the two hyperplanes. From this we conclude that $\|p_A(x) - p_A(y)\| \leq \|x - y\|$. □

The above result proves that the projection $p_A(x)$ is indeed a contraction. Since contraction mappings do not increase lengths, we can use this fact to demonstrate that the boundary of A_r is less than that of A.

Theorem 12. *If the set $A \subset \mathbb{R}^D$ is compact and convex, then $\Phi(A_r) = |\partial A_r| \leq |\partial A| = \Phi(A)$.*

Proof. Consider the projection $p_{A_r} : A \to A_r$. Note that for $r > 0$, we have that $A_r = \{x \in A \mid d(x, A^c) \geq r\}$ is entirely contained in the interior of A, which implies that $A_r \cap \partial A = \emptyset$. Thus, our situation satisfies the assumption we made in the proof of Lemma 25.

Under the action of p_{A_r}, any point in $\mathbb{R}^D$ outside of A_r will get mapped to ∂A_r. In

particular, since the map is onto, ∂A will get mapped to ∂A_r, i.e. $p(\partial A) = \partial A_r$. Using the fact that p_{A_r} is contractive, we have that

$$\Phi(A_r) = |\partial A_r| = |p(\partial A)| \leq |\partial A| = \Phi(A),$$

from which we conclude that A is a shrinkable set. □

This provides the final step in our proof of Lemma 22. Note that we do not require an explicit formula for the surface area of the boundary of a D-dimensional convex set. In general, one may use the Cauchy-Crofton formula to calculate this quantity—for more details, see Theorem 5.5.2 of Ref. [202].

Appendix C: Supplemental Material for Sec. 6.2

C.1 Proof of Lemma 5

In this section we provide a proof of Lemma 5. One direction of the lemma is straightforward. If the joint state is a product, i.e. $\rho = \rho_{\mathcal{X}} \otimes \rho_{\tilde{\mathcal{X}}}$, then all bipartite disconnected correlators between $A_{\mathcal{X}} \in \mathcal{S}(\mathcal{X})$ and $A_{\tilde{\mathcal{X}}} \in \mathcal{S}(\tilde{\mathcal{X}})$ are factorizable, $\langle A_{\mathcal{X}} A_{\tilde{\mathcal{X}}} \rangle = \langle A_{\mathcal{X}} \rangle \langle A_{\tilde{\mathcal{X}}} \rangle$. Therefore all bipartite connected correlators vanish. To prove the opposite direction, that is vanishing of all bipartite connected correlators implies ρ is a product state, let $\left\{\Gamma_\mu^{\mathcal{X}}\right\}$ denote a complete normalized basis for density matrices of $\mathcal{X}$, and likewise for $\left\{\Gamma_\mu^{\tilde{\mathcal{X}}}\right\}$. Any joint state of $\mathcal{X}$ and $\tilde{\mathcal{X}}$ may be written as

$$\rho = \frac{1}{N}\Bigg(\mathbb{I}_{\mathcal{X}\cup\tilde{\mathcal{X}}} + \sum_\mu \left\langle \Gamma_\mu^{\mathcal{X}} \right\rangle \Gamma_\mu^{\mathcal{X}} \otimes \mathbb{I}_{\tilde{\mathcal{X}}} + \sum_\nu \left\langle \Gamma_\nu^{\tilde{\mathcal{X}}} \right\rangle \mathbb{I}_{\mathcal{X}} \otimes \Gamma_\nu^{\tilde{\mathcal{X}}} + \sum_{\mu,\nu} \left\langle \Gamma_\mu^{\mathcal{X}} \otimes \Gamma_\nu^{\tilde{\mathcal{X}}} \right\rangle \Gamma_\mu^{\mathcal{X}} \otimes \Gamma_\nu^{\tilde{\mathcal{X}}}\Bigg), \tag{C.1}$$

where $N = |\mathcal{H}_{\mathcal{X}} \otimes \mathcal{H}_{\tilde{\mathcal{X}}}|$ is the dimension of the joint Hilbert space. Since all bipartite connected correlators vanish,

$$\left\langle \Gamma_\mu^{\mathcal{X}} \otimes \Gamma_\nu^{\tilde{\mathcal{X}}} \right\rangle = \left\langle \Gamma_\mu^{\mathcal{X}} \right\rangle \left\langle \Gamma_\nu^{\tilde{\mathcal{X}}} \right\rangle \tag{C.2}$$

for all μ, ν. Therefore ρ is also factorizable,

$$\rho = \frac{1}{N}\left(\mathbb{I}_{\mathcal{X}} + \sum_{\mu} \left\langle \Gamma_{\mu}^{\mathcal{X}} \right\rangle \Gamma_{\mu}^{\mathcal{X}}\right) \otimes \left(\mathbb{I}_{\tilde{\mathcal{X}}} + \sum_{\nu} \left\langle \Gamma_{\nu}^{\tilde{\mathcal{X}}} \right\rangle \Gamma_{\nu}^{\tilde{\mathcal{X}}}\right). \tag{C.3}$$

Thus the lemma follows.

C.2 Equivalent Definitions of Multipartite Connected Correlators

In this section we present some definitions of the multipartite connected correlation function which are equivalent to Eq. (6.14). The multipartite connected correlator can also be generated by [127]:

$$u_n(A_1, \ldots, A_n) = \left[\frac{\partial^n}{\partial\lambda_1 \ldots \partial\lambda_n} \ln \left\langle e^{\sum_{i=1}^n \lambda_i A_i} \right\rangle\right]_{\vec{\lambda}=0}, \tag{C.4}$$

where the partial derivative is evaluated at $\vec{\lambda} = (\lambda_1, \ldots, \lambda_n) = 0$. This generating form will be used in Appendix C.5 to evaluate multipartite connected correlators of the GHZ state. An equivalent way to define multipartite connected correlators is via lower-order correlators,

$$u_n\left(A_1, \ldots, A_n\right) = \langle A_1 \ldots A_n \rangle - \sum_{P}{}' \prod_{p \in P} u_{|p|}\left(\tilde{A}_p\right), \tag{C.5}$$

where the sum $\sum_P'$ is taken over all partitions of $\{\mathcal{X}_1, \ldots, \mathcal{X}_n\}$ except for the trivial partition $P = \{\mathcal{X}_1, \ldots, \mathcal{X}_n\}$, and $\tilde{A}_p = \{A_i : i \in p\}$ denotes the set of all observables with indices in set p. We shall find this definition useful for the inductive proof of Theorem 9

and in Appendix C.5.

C.3 Proof of Lemma 6

In this section we prove the connection between factorizability and vanishing connected correlators in Lemma 6. We shall prove this lemma inductively using generating functions of multipartite connected correlators (C.4),

$$\begin{aligned}&\ln\left\langle \exp\left(\sum_{i=1}^{k_1}\lambda_i A_i+\sum_{j=1}^{k_2}\lambda'_j B_j\right)\right\rangle\\&=\ln\left\langle \exp\left(\sum_{i=1}^{k_1}\lambda_i A_i\right)\right\rangle+\ln\left\langle \exp\left(\sum_{j=1}^{k_2}\lambda'_j B_j\right)\right\rangle.\end{aligned}\tag{C.6}$$

The first term on the right hand side is independent of any λ'_j. Therefore, partial derivatives with respect to λ'_js will make the first term vanish. Similarly, the second term will also vanish after partial derivatives with respect to λ_is. Therefore multipartite connected correlators, which are n^{th} order partial derivatives of the left hand side with respect to both λ_is and λ'_js, will vanish. The lemma follows.

C.4 Proof of Theorem 9

In this section we prove Theorem 9 by induction on n. When $n=2$, the inequalities reduce to bipartite Lieb-Robinson bounds. Assuming that it holds for any $2\leq n\leq k-1$, we shall prove that it holds for $n=k$. We start with the recursive definition of connected

correlators (Appendix C.2):

$$\langle A_1 \dots A_k \rangle = \sum_{P \in \mathcal{P}(S)} \prod_{p \in P} u_{|p|}\left(\tilde{A}_p\right), \tag{C.7}$$

where $\mathcal{P}(S)$ denotes the set of all partitions of $S = 1, \dots, k$, and $\tilde{A}_p = \{A_i : i \in p\}$ denotes the set of all observables with indices in set p. Consider one particular bipartition of S, e.g. $S = S_1 \cup S_2$ such that $S_1 \cap S_2 = \varnothing$. The partitions of S can then be divided into two types. Partitions of the first type have elements that lie entirely on either S_1 or S_2. They therefore belong to the set $\mathcal{P}(S_1) \oplus \mathcal{P}(S_2)$. The sum over these partitions in Eq. (C.7) can then be factored into a product of two sums over $\mathcal{P}(S_1)$ and $\mathcal{P}(S_2)$,

$$\begin{aligned} &\left[\sum_{P_1 \in \mathcal{P}(S_1)} \prod_{p_1 \in P_1} u_{|p_1|}\left(\tilde{A}_{p_1}\right)\right]\left[\sum_{P_2 \in \mathcal{P}(S_2)} \prod_{p_2 \in P_2} u_{|p_2|}\left(\tilde{A}_{p_2}\right)\right] \\ &= \left\langle \prod_{i \in S_1} A_i \right\rangle \left\langle \prod_{i \in S_2} A_i \right\rangle, \end{aligned} \tag{C.8}$$

where we have used the definition (C.7) for the sets S_1 and S_2. The terms in Eq. (C.7) we have not yet summed over are partitions in which some elements overlap with both S_1 and S_2, namely $\mathcal{P}(S) \setminus \mathcal{P}(S_1) \oplus \mathcal{P}(S_2) \equiv \mathcal{P}_{12}$. We can then rewrite Eq. (C.7) as

$$\begin{aligned} \langle A_1 \dots A_k \rangle =& u_k(A_1, \dots, A_k) + \left\langle \prod_{i \in S_1} A_i \right\rangle \left\langle \prod_{i \in S_2} A_i \right\rangle \\ &+ \sum_{P_3 \in \mathcal{P}_{12}} \prod_{p_3 \in P_3} u_{|p_3|}\left(\tilde{A}_{p_3}\right). \end{aligned} \tag{C.9}$$

Rearranging Eq. (C.9) in terms of bipartite connected correlators, we have

$$u_k(A_1,\ldots,A_k) = u_2\left(\prod_{i\in S_1} A_i, \prod_{i\in S_2} A_i\right) - \sum_{P_3\in\mathcal{P}_{12}}\prod_{p_3\in P_3} u_{|p_3|}\Big(\{A_{i\in p_3}\}\Big). \tag{C.10}$$

Therefore,

$$|u_k(A_1\ldots A_k)| \leq \left|u_2\left(\prod_{i\in S_1} A_i, \prod_{i\in S_2} A_i\right)\right| + \sum_{P_3\in\mathcal{P}_{12}}\prod_{p_3\in P_3}\left|u_{|p_3|}\Big(\{A_{i\in p_3}\}\Big)\right|. \tag{C.11}$$

The first term is bounded by $\propto \exp(vt - d(S_1, S_2))$, where the distance between subsystems S_1 and S_2, i.e. $d(S_1, S_2)$, is defined as the smallest separation distance between a site in S_1 and a site in S_2. To bound the second term, we first realize that the connected correlators here are between at most $k-1$ points, and therefore our induction hypothesis applies. For each connected correlator u, there can be two possibilities. It can involve subsystems supported by both S_1 and S_2, or supported by either S_1 or S_2 alone. If we sum over those of the second type, we again get expectation values which are bounded by 1. For the connected correlator u that involves qubits in both S_1 and S_2, by the induction hypothesis it is bounded by $\exp(vt-r)$, where r is the largest distance between any bipartitions of the subsystems. By dividing those subsystems into those in S_1 and those in S_2, the distance r has to be at least the one between S_1 and S_2, i.e. $r \geq d(S_1, S_2)$. Therefore

the second term in Eq. (C.11) is also bounded by $\exp(vt - d(S_1, S_2))$. In the end, we get

$$|u_k(A_1, \ldots, A_k)| \leq C_k \exp\left[v_{\mathrm{LR}} t - d(S_1, S_2)\right] \tag{C.12}$$

for some constant C_k to be determined. For each choice of bipartition $\{S_1, S_2\}$, we get one such inequality. The tightest bound is obtained from the bipartition with the largest distance d, i.e.

$$|u_k(A_1, \ldots, A_k)| \leq C_k \exp\left[v_{\mathrm{LR}} t - R\right] \tag{C.13}$$

with $R = \max_{S_1} d(S_1, S_2)$. Thus the hypothesis is true for $n = k$, and by induction it holds for any n.

We now prove the second part of the theorem, i.e. $C_n \leq n^n \frac{C_2}{4}$. Clearly it holds for $n = 2$. We prove that if the statement holds up to $n = k - 1$, it must also hold for $n = k$. Recall that a k-point connected correlator is bounded by Eq. (C.11). The first term of Eq. (C.11) is bounded by 1. We need to find a bound for the sum. Note that at the critical time $t = R/v$, the only non-negligible contributing terms are those involving $S_1' \subset S_1$ and $S_2' \subset S_2$ such that the distance between S_1' and S_2' is exactly R (by construction the distance is at least R).

Let $S_1^{(0)} \subset S_1$ and $S_2^{(0)} \subset S_2$ be such that the distance between any $s_1 \in S_1^{(0)}$ and $s_2 \in S_1^{(0)}$ is always R. The point is that only connected correlators that involve such s_1 and s_2 will contribute to the sum. We now count the contribution from such correlators. If we take k_1 subsystems from $S_1^{(0)}$, k_2 subsystems from $S_2^{(0)}$ and k_3 subsystems from

$S_3^{(0)} = S \backslash S_1^{(0)} \cup S_2^{(0)}$, their contribution is $\mathcal{O}\left((k_1+k_2+k_3)^{k_1+k_2+k_3}\right)$. Note that summing over connected correlators of leftover subsystems, we get their disconnected correlator, which is bounded by 1. Note also that by counting this way, some terms will appear more than once, so we get a loose bound. Denoting by m_1, m_2, m_3 the size of $S_1^{(0)}, S_2^{(0)}$ and $S_3^{(0)}$, we can bound the constant C_k by summing over all possible choices of $k_1+k_2+k_3 \leq k-1$,

$$C_k \leq \frac{C_2}{4} \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \sum_{k_3=0}^{m_3} \binom{m_1}{k_1} \binom{m_2}{k_2} \binom{m_3}{k_3} (k_1+k_2+k_3)^{k_1+k_2+k_3} \tag{C.14}$$

$$\leq \frac{C_2}{4} \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \sum_{k_3=0}^{m_3} \binom{m_1}{k_1} \binom{m_2}{k_2} \binom{m_3}{k_3} (k-1)^{k_1+k_2+k_3} \tag{C.15}$$

$$= \frac{C_2}{4} (k^{m_1}-1)(k^{m_2}-1)k^{m_3} < \frac{C_2}{4} k^{m_1+m_2+m_3} = k^k \frac{C_2}{4}. \tag{C.16}$$

Thus $C_k \leq k^k \frac{C_2}{4}$ holds for $n = k$, and by induction it holds for any n.

C.5 Calculation of Connected Correlators

In this section we show how connected correlators are calculated for the GHZ states, the cluster states and the product state evolved under the XX Hamiltonian.

C.5.1 The GHZ State

The generating function of $u_n(Z_1, \ldots, Z_n)$ evaluated for the GHZ state of n qubits is

$$g_n \equiv \ln \left\langle \exp\left\{ \sum_{i=1}^{n} \lambda_i Z_i \right\} \right\rangle_{\text{GHZ}} \tag{C.17}$$

$$= \ln \left[\frac{1}{2} \exp\left(\sum_{i=1}^{n} \lambda_i \right) + \frac{1}{2} \exp\left(-\sum_{i=1}^{n} \lambda_i \right) \right] \tag{C.18}$$

$$= \ln \left[\cosh\left(\sum_{i=1}^{n} \lambda_i \right) \right]. \tag{C.19}$$

Let $\lambda \equiv \sum_{i=1}^{n} \lambda_i$. Then

$$\frac{\partial g_n}{\partial \lambda_i} = \frac{\partial g_n}{\partial \lambda} \frac{\partial \lambda}{\partial \lambda_i} = \frac{\partial g_n}{\partial \lambda} \tag{C.20}$$

for all i. Therefore the multipartite connected correlator above is given by

$$u_n(Z_1, \ldots, Z_n) = \left. \frac{\partial^n g_n}{\partial \lambda^n} \right|_{\lambda=0} = \left[\frac{\partial^n}{\partial \lambda^n} \ln(\cosh \lambda) \right]_{\lambda=0}. \tag{C.21}$$

Note that this connected correlator has the same parity as n. Therefore for odd n, it vanishes. For even n, the correlator is given by

$$u_n = \frac{2^n (2^n - 1) B_n}{n}, \tag{C.22}$$

where B_n is the n^{th} Bernoulli number. In the large n limit, the Bernoulli number is approximated by

$$|B_n| \approx 4\sqrt{\frac{\pi n}{2}}\left(\frac{n}{2\pi e}\right)^n. \tag{C.23}$$

Therefore the n-point connected correlator of the GHZ state grows as $u_n \propto n^{-1/2}(\frac{2}{\pi e})^n n^n = \mathcal{O}(n^n)$.

C.5.2 The Cluster States

For each vertex i in a cluster state's graph, we can associate an operator $X_i \prod_{j\in\mathcal{N}(i)} Z_j$, where $\mathcal{N}(i)$ denotes the set of vertices adjacent to i. These operators generate a stabilizer group of which the cluster state is a simultaneous eigenstate. Operators outside of this group have no disconnected correlations. Using the stabilizer group, we can count the number of contributing disconnected correlators in the definition of connected correlators (6.14). For example, for the observables $Y_1, X_2, X_3, \ldots, X_{n-1}, Y_n$ in the cluster state in Fig. 6.3(a), all low-order disconnected correlators vanish. Therefore,

$$\begin{aligned} &u_n(Y_1, X_2, X_3, \ldots, X_{n-1}, Y_n) \\ &= \langle Y_1 X_2 X_3 \ldots X_{n-1} Y_n \rangle = 1. \end{aligned} \tag{C.24}$$

Similarly, by direct counting we find the n-point connected correlator of the Fig. 6.3(b) cluster state $u_n(\{T_j : j = 1, \ldots, n\}) = 2^{\frac{n-1}{3}}$, where $T_j = X_j$ for all $1 < j < n$ such that $j \equiv 1 \pmod 3$, and $T_j = Y_j$ otherwise.

C.5.3 A Product State Evolved Under the XX Hamiltonian

The time evolution shown in Fig. 6.4 can be verified as follows. The time-dependent state of n qubits evolving from $|00\ldots0\rangle$ under $H = \sum_{\langle i,j\rangle} X_i X_j$ can be written in the form of a matrix product state,

$$|\psi(t)\rangle = \sum_{i_1,\ldots,i_n\in\{0,1\}} c_{i_1 i_2 \ldots i_n}(t)\,|i_1 i_2 \ldots i_n\rangle\,, \tag{C.25}$$

the coefficients of which are given by

$$c_{i_1 i_2 \ldots i_n}(t) = L_{i_1} A_{i_2}(t) A_{i_3}(t) \cdots A_{i_{n-1}}(t) R_{i_n}(t), \tag{C.26}$$

where

$$L_0 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \tag{C.27}$$

$$L_1 = \begin{pmatrix} 0 & 1 \end{pmatrix}, \tag{C.28}$$

$$A_0(t) = \begin{pmatrix} \cos t & 0 \\ 0 & -i \sin t \end{pmatrix}, \tag{C.29}$$

$$A_1(t) = \begin{pmatrix} 0 & \cos t \\ -i \sin t & 0 \end{pmatrix}, \tag{C.30}$$

$$R_0(t) = \begin{pmatrix} \cos t \\ 0 \end{pmatrix}, \tag{C.31}$$

$$R_1(t) = \begin{pmatrix} 0 \\ -i \sin t \end{pmatrix}. \tag{C.32}$$

Note that this matrix product state is in left canonical form (i.e. $\sum_i L_i^\dagger L_i = \sum_i A_i^\dagger A_i = I$) and it is normalized ($\sum_i R_i^\dagger R_i = 1$). Our goal is to first determine all disconnected correlators of the form $\langle O_1 O_2 \cdots O_n \rangle$ where O_i is either I or Z. Because all such operators are diagonal on each site, we can write the expectation value itself as a matrix product. In the end, we find that the disconnected correlator picks up a factor of $\cos(2t)$ for each "boundary" between a Z operator and an I operator. For instance, on a 5-qubit system, the expectation value $\langle Z_2 Z_3 Z_5 \rangle = \langle IZZIZ \rangle = [\cos(2t)]^3$, as there are 3 relevant boundaries: between qubits 1–2, 3–4, and 4–5.

From this, it is already obvious that our connected correlator $u_n(Z_1, \ldots, Z_n)$ will

be some polynomial of the variable $\cos(2t)$. Given some partition $\mathcal{P}$, we would like to determine the power to which $\cos(2t)$ is raised. Let us, for sake of example, denote our partition by letters of the alphabet. On 5 qubits, ABBCA corresponds to the product of disconnected correlators $\langle Z_1 Z_5 \rangle \langle Z_2 Z_3 \rangle \langle Z_4 \rangle = \langle ZIIIZ \rangle \langle IZZII \rangle \langle IIIZI \rangle = [\cos(2t)]^6$. In general, the product of disconnected correlators will be $[\cos(2t)]^{2v}$ where v is the number of bonds that border two distinct subsets of the partition. (In the case of the example ABBCA, this includes each bond except the one between sites 2–3, which are both in the same subset, B.)

Now we would like to count the number of partitions which contribute to the term with power $2v$. Because the coefficient in the connected correlator depends on the number of subsets in the partition $|\mathcal{P}|$, we must consider separately partitions with different numbers of subsets. Given n qubits, there are $n-1$ bonds between qubits. Thus there are $\binom{n-1}{v}$ different ways to choose v bonds which connect different subsets of the partition. Given these v bonds, there are $\left\{ {v \atop a} \right\}$ different ways to construct partitions with $(a+1)$ total subsets. (Here, $\left\{ {v \atop a} \right\}$ denotes a Stirling number of the second kind.) Thus, the number of partitions on n qubits with v bonds that border two distinct subsets and with $(a+1)$ total subsets is $\binom{n-1}{v} \left\{ {v \atop a} \right\}$. Note that $\sum_{v=0}^{n-1} \binom{n-1}{v} \sum_{a=0}^{v} \left\{ {v \atop a} \right\}$ is equal to the n^{th} Bell number $\mathcal{B}_n$, so we have indeed accounted for all possible partitions.

As mentioned above, given a partition, two factors of $\cos(2t)$ are picked up for each bond that borders two distinct subsets. In general, we can compute the expectation value of the connected correlator from Eq. (6.14) as follows:

$$
\begin{aligned}
&u_n\left(Z_1,\ldots,Z_n\right)\\
&=\sum_{\mathcal{P}}(-)^{|\mathcal{P}|-1}(|\mathcal{P}|-1)!\prod_{P\in\mathcal{P}}\left\langle\prod_{p\in P}Z_p\right\rangle=\sum_{v=0}^{n-1}\sum_{a=0}^{v}(-1)^a\,a!\binom{n-1}{v}\left\{{v\atop a}\right\}\left[\cos(2t)\right]^{2v}\\
&=\sum_{v=0}^{n-1}\binom{n-1}{v}\left[\cos(2t)\right]^{2v}\sum_{a=0}^{v}(-1)^a\,a!\left\{{v\atop a}\right\}=\sum_{v=0}^{n-1}\binom{n-1}{v}\left[\cos(2t)\right]^{2v}(-)^v\\
&=\sum_{v=0}^{n-1}\binom{n-1}{v}\left[-\cos^2(2t)\right]^v=\left[1-\cos^2(2t)\right]^{n-1}=\left[\sin^2(2t)\right]^{n-1},
\end{aligned}
\tag{C.33}
$$

where we have used the identity $\sum_{a=0}^{v}(-1)^a\,a!\left\{{v\atop a}\right\}=(-1)^v$ [203].

Appendix D: Supplemental Material for Sec. 6.9

D.1 The Generalization of Gong et al.to Many-Body Interactions

In this section, we prove Eq. (6.83) and thereby generalize the bound in Gong et al. [29] from two-body to k-body interactions, where k is an arbitrary finite integer. This bound is an ingredient in the generalization of the tighter Lieb-Robinson bound in Tran et al. [21] to k-body interactions.

Proof. We recall that the bound in Ref. [29] is based on the Hastings & Koma series [18]:

$$\|[A(t),B]\| \leq 2\|A\|\|B\| \sum_{k=1}^{\infty} \frac{(2t)^k}{k!} \left[\sum_{Z_1 : Z_1 \cap X \neq \varnothing} \sum_{Z_2 : Z_1 \cap Z_2 \neq \varnothing} \cdots \sum_{\substack{Z_k : Z_{k-1} \cap Z_k \neq \varnothing, \\ Z_k \cap Y \neq \varnothing}} \prod_{i=1}^{k} \|h_{Z_i}\| \right], \tag{D.1}$$

and we can bound the summation within the square brackets as

$$\begin{aligned} &\sum_{Z_1 : Z_1 \cap X \neq \varnothing} \sum_{Z_2 : Z_1 \cap Z_2 \neq \varnothing} \cdots \sum_{\substack{Z_k : Z_{k-1} \cap Z_k \neq \varnothing, \\ Z_k \cap Y \neq \varnothing}} \prod_{i=1}^{k} \|h_{Z_i}\| \\ &\leq \sum_{i \in X} \sum_{j \in Y} \sum_{z_1} \sum_{z_2} \cdots \sum_{z_{k-1}} \left(\sum_{Z_1 \ni i, z_1} \|h_{Z_1}\| \right) \cdots \left(\sum_{Z_k \ni z_{k-1}, j} \|h_{Z_k}\| \right) \\ &\leq \sum_{i \in X} \sum_{j \in Y} \lambda^k \mathcal{J}^k(i,j), \end{aligned} \tag{D.2}$$

where $\mathcal{J}^k(i,j)$ is given by the k-fold convolution of the hopping terms $J_{ij} \equiv \frac{1}{r_{ij}^\alpha}$ (where $r_{ij} = \text{dist}\,(()\,i,j)$) for $i \neq j$ and $J_{ii} = 1$ for all i:

$$\mathcal{J}^k(i,j) \equiv \sum_{z_1} \sum_{z_2} \cdots \sum_{z_{k-1}} J_{iz_1} J_{z_1 z_2} \ldots J_{z_{k-1},j}.$$

Note that Eq. (D.2) comes from Definition 1: $\sum_{Z \ni i,j} \|h_Z\| \leq 1/r_{ij}^\alpha = J_{ij}$ for $i \neq j$ and

$$\sum_{Z \ni i} \|h_Z\| \leq \sum_j \sum_{Z \ni i,j} \|h_Z\| \leq \lambda, \tag{D.3}$$

where $\lambda = 1 + \sum_{j \neq i} 1/r_{ij}^\alpha$ is a finite constant for all $\alpha > d$. This equation is exactly Eq. (3) in Ref. [29].

For simplicity, we consider $d = 1$ in the following discussion. To put a bound on $\mathcal{J}^k(i,j)$, we use the same trick as in Ref. [29]. First, we consider the sum over z_1:

$$\sum_{z_1} J_{iz_1} J_{z_1 z_2} \leq 2 \sum_{z_1 : r_{iz_1} \leq r_{z_1 j}} J_{iz_1} J_{z_1 z_2}, \tag{D.4}$$

where the right hand side sums only over z_1 being closer to i than to z_2 and the factor 2 accounts for exchanging the roles of i and z_2. We further separate the sum over z_1 in Eq. (D.3) into two, corresponding to whether z_1 is within a unit distance from i or not:

$$\sum_{z_1} J_{iz_1} J_{z_1 z_2} \leq 2 \left(\sum_{z_1 : r_{iz_1} \leq 1} + \sum_{z_1 : r_{iz_1} \geq 2} \right) J_{iz_1} J_{z_1 z_2}. \tag{D.5}$$

Since $r_{iz_1} \leq r_{z_1 z_2}$, it follows that $r_{z_1 z_2} \geq r_{iz_2}/2$. Therefore, $J_{z_1 z_2} \leq 2^\alpha J_{iz_2}$ and we further

bound the second sum in Eq. (D.5) by

$$\begin{aligned}\sum_{z_1:r_{iz_1}\geq 2} J_{iz_1}J_{z_1z_2} &\leq 2^{\alpha}J_{iz_2}\sum_{z_1:r_{iz_1}\geq 2} J_{iz_1}\\ &\leq 2^{\alpha}J_{iz_2}2^{1-\alpha}(\lambda-1)\\ &\leq 2(\lambda-1)\sum_{z_1:r_{iz_1}\leq 1} J_{iz_1}J_{z_1z_2}, \end{aligned} \tag{D.6}$$

where we bound $\sum_{z_1:r_{iz_1}\geq 2} J_{iz_1} \leq 2^{1-\alpha}(\lambda-1)$ and $J_{iz_2} \leq \sum_{z_1:r_{iz_1}\leq 1} J_{iz_1}J_{z_1z_2}$ similarly to Ref. [29]. Therefore, we have $\sum_{z_1} J_{iz_1}J_{z_1z_2} \leq 4\lambda \sum_{z_1:r_{iz_1}\leq 1} J_{iz_1}J_{z_1z_2}$. Repeating this analysis for $z_2,\ldots,z_k$ in Eq. (D.3), we have an upper bound on $\mathcal{J}^k(i,j)$:

$$\begin{aligned}\mathcal{J}^k(i,j) &\leq (4\lambda)^{k-1}\sum_{z_1:r_{iz_1}\leq 1}\sum_{z_2:r_{z_1z_2}\leq 1}\cdots\\ &\quad\cdots\sum_{z_{k-1}:r_{z_{k-2}z_{k-1}}\leq 1} J_{iz_1}J_{z_1z_2}\ldots J_{z_{k-1},j} \qquad \text{(D.7)}\\ &\leq (12\lambda)^{k-1}\times\begin{cases}1/(r_{ij}-k+1)^{\alpha} & \text{if } k<\mu r_{ij},\\ 1 & \text{if } k\geq \mu r_{ij},\end{cases} \qquad \text{(D.8)}\\ &\leq (12\lambda)^{k-1}\times\begin{cases}1/[(1-\mu)r_{ij}]^{\alpha} & \text{if } k<\mu r_{ij},\\ 1 & \text{if } k\geq \mu r_{ij},\end{cases} \qquad \text{(D.9)}\end{aligned}$$

where $\mu\in(0,1)$ is an arbitrary constant.

To get the second to last bound, we note that the maximum value that the summand in Eq. (D.7) may achieve is $1/(r_{ij}-k+1)^{\alpha}$ when $k<\mu r_{ij}$ and 1 when $k\geq r_{ij}$, and the number of sites within a unit distance of any site is 3. Plugging this bound into Eq. (D.1)

and Eq. (D.2), we have the Lieb-Robinson bound in Ref. [29] generalized to many-body interactions:

$$
\begin{aligned}
&\|[A(t), B]\| \\
&\leq \|A\|\|B\| \sum_{i\in X}\sum_{j\in Y}\left(\sum_{k=1}^{\lceil \mu r_{ij}\rceil -1} \frac{(24\lambda^2 t)^k}{6\lambda k![(1-\mu)r_{ij}]^\alpha} + \sum_{k=\lceil \mu r_{ij}\rceil}^{\infty} \frac{(24\lambda^2 t)^k}{6\lambda k!}\right) \\
&\leq \|A\|\|B\| \sum_{i\in X}\sum_{j\in Y} Ce^{vt}\left(\frac{1}{[(1-\mu)r_{ij}]^\alpha} + e^{-\mu r_{ij}}\right) \qquad \text{(D.10)} \\
&\leq \|A\|\|B\||X||Y|Ce^{vt}\left(\frac{1}{[(1-\mu)r]^\alpha} + e^{-\mu r}\right), \qquad \text{(D.11)}
\end{aligned}
$$

where $C = 1/6\lambda$, $v = 24\lambda^2$, and r is, again, the distance between X, Y. The proof for $d > 1$ follows a very similar analysis. □

A feature of Eq. (D.11) is that it depends on $|X|, |Y|$, which can become problematic when A, B are supported on a large number of sites. In such cases, we can sum over the sites of X, Y in Eq. (D.10) to get more useful bounds. Without any other assumptions, summing over the sites of Y gives an extra factor of r^d:

$$
\|[A(t),B]\| \leq C\|A\|\|B\||X|\left(\frac{1}{(1-\mu)^\alpha r^{\alpha-d}} + \frac{e^{vt}}{e^{vt-\mu r}}\right), \qquad \text{(D.12)}
$$

where the constant C absorbs all constants that may depend on μ. Note that the bound still depends on $|X|$ but not on $|Y|$.

We can go one step further and sum over the sites of X, but we need to assume that

X is convex (similarly to Ref. [21]). Then, we have a bound

$$\|[A(t),B]\| \leq C\|A\|\|B\|\phi(X)\left(\frac{1}{(1-\mu)^{\alpha}r^{\alpha-d-1}}+e^{vt-\mu r}\right), \tag{D.13}$$

which is independent of $|X|$. Here $\phi(X)$ is the boundary area of X, defined as the number of sites in X that are adjacent to a site outside X.

D.2 The Absorption Rate from Linear Response Theory

This section provides more details on the derivation of the absorption rate within linear response theory. In particular, we provide more mathematically rigorous proofs of Eq. (6.106) [Appendix D.2.1] and Eq. (6.109) [Appendix D.2.2].

D.2.1 Proof of Eq. (6.106)

In this section, we prove the statement of Eq. (6.106) [also Eq. (D.15) below]. We recall that the system Hamiltonian H_0 is a power-law Hamiltonian, while the harmonic drive $V(t) = g\cos(\omega t)O$ is a sum of local terms, $g\cos(\omega t)O_i$, each of which is supported on the site i only, where i runs over the sites of the system. In addition, we assume that the system is initially in the equilibrium state ρ_β of H_0 corresponding to the temperature $1/\beta$. To the lowest order in g, the energy absorption rate of the system is proportional to the dissipative (imaginary) part of the response function, $\sigma(\omega) = \sum_{i,j}\sigma_{ij}(\omega)$, where i, j

are the sites of the system and

$$\sigma_{ij}(\omega) = \frac{1}{2}\int_{-\infty}^{\infty} dt e^{i\omega t} \left\langle [O_i(t), O_j(0)] \right\rangle_\beta , \tag{D.14}$$

where $\langle X \rangle_\beta \equiv \mathrm{Tr}(\rho_\beta X)$ denotes the expectation value of an operator X in ρ_β.

In Ref. [146], the authors proved that there exists constants C, κ such that for all $\omega > 0, \delta\omega > 0$ and for all pairs i, j,

$$|\sigma_{ij}([\omega, \omega + \delta\omega])| \leq C e^{-\kappa\omega}. \tag{D.15}$$

The statement in Ref. [146] is for finite-range interactions, but, for completeness, we show here that it also holds for power-law Hamiltonians. First, we consider the diagonal terms $\sigma_{ii}(\omega)$. Let $|n\rangle$ and E_n denote the eigenstates and eigenvalues of H_0. Similarly to Ref. [146], we rewrite $\sigma_{ii}(\omega)$ as

$$\sigma_{ii}(\omega) = \pi \sum_n p_n [\gamma_{ii}^{(n)}(\omega) - \gamma_{ii}^{(n)}(-\omega)], \tag{D.16}$$

where p_n is the probability that the state is in the eigenstate $|n\rangle$, and $\gamma_{ii}^{(n)}$ denotes the contribution to σ_{ii} from the n-th eigenstate:

$$\begin{aligned} \gamma_{ii}^{(n)}(\omega) &= \sum_m |\left\langle m \right| O_i \left| n \right\rangle|^2 \delta(E_n - E_m - \omega) \\ &= \sum_m \frac{|\left\langle m \right| \mathrm{ad}_H^k O_i \left| n \right\rangle|^2}{\omega^{2k}} \delta(E_n - E_m - \omega), \end{aligned} \tag{D.17}$$

where $\mathrm{ad}_H O_i = [H, O_i]$, k is an integer to be chosen later, and the last equality comes

from the fact that $|m\rangle, |n\rangle$ are eigenstates of H and the δ function fixes the energy difference to be ω.

In Ref. [146], the authors used the fact that H has a finite range to upper bound the norm of $\text{ad}_H^k O_i$ by $\lambda^k k!$ for some constant λ. For power-law interactions, the proof does not apply because the Hamiltonian H can contain interaction terms between arbitrarily far sites. Instead, we upper bound $\text{ad}_H^k O_i$ by realizing that O_i technically satisfies Definition 1 and is therefore a power-law Hamiltonian. It then follows from Lemma 27 in Appendix D.5 that $\text{ad}_H^k O_i \in \lambda^k k! \mathbb{H}_\alpha$, i.e. $\text{ad}_H^k O_i$ is a power-law Hamiltonian up to a factor $\lambda^k k!$, where λ is the same constant as in Lemma 27 and $\mathbb{H}_\alpha$ is the set of power-law Hamiltonians with exponent α [See Appendix D.5.1]. Finally, we can upper bound

$$\|\text{ad}_H^k O_i\| \leq C\lambda^k k!, \tag{D.18}$$

by realizing that the supports of the terms in $\text{ad}_H^k O_i$ all contain the site i.

Integrating Eq. (D.17) over ω, assuming $\delta\omega$ is small enough so that the number of energy levels in the range $[\omega, \omega + \delta\omega]$ is finite, and using Eq. (D.18), we have

$$\begin{aligned} |\gamma_{ii}^{(n)}([\omega, \omega + \delta\omega])| &\leq C \left(\frac{\lambda^k k!}{\omega^k}\right)^2 \\ &\leq C \left(\frac{\lambda k}{\omega}\right)^{2k} \leq C e^{-\kappa\omega}, \end{aligned} \tag{D.19}$$

where $\kappa = 2/(\lambda e)$ and, to get the last line, we choose $k = \omega/(\lambda e)$. Plugging this bound into Eq. (D.16) and summing over n yields Eq. (6.106) for $i = j$. The bound for $i \neq j$

can be derived using the positivity of σ [146] and the Cauchy-Schwartz inequality,

$$|\sigma_{ij}(\omega)| \leq \frac{1}{2}[\sigma_{ii}(\omega) + \sigma_{jj}(\omega)]. \tag{D.20}$$

Therefore, Eq. (6.106) applies for all power-law Hamiltonians H.

D.2.2 Proof of Eq. (6.109)

We now provide a rigorous proof of Eq. (6.109) in the main text. Equation (D.15) says that the (i, j) entry of $\sigma([\omega, \omega + \delta\omega])$ is exponentially suppressed. In principle, summing over all i, j implies that $\sigma([\omega, \omega + \delta\omega])$ is also exponentially small as a function of ω. However, since there are N sites in the system, this summation results in an additional factor of N^2, making $\sigma([\omega, \omega + \delta\omega])$ superextensive. Therefore, this naive calculation breaks down in the thermodynamic limit ($N \to \infty$).

Instead, to show that $\sigma([\omega, \omega + \delta\omega])$ increases only as fast as N, we use Lieb-Robinson bounds to bound the off diagonal terms $\sigma_{ij}(\omega)$. Let $r_{ij} = \text{dist}\,(()\, i, j)$ denote the distance between the pair of sites i, j. Without loss of generality, we assume $\omega \geq 2\delta\omega$. We can then bound

$$\begin{aligned}
&\sigma([\omega, \omega + \delta\omega]) = \int_{\omega}^{\omega+\delta\omega} d\omega' \sigma(\omega') \\
&\leq c_1 \int_{-\infty}^{\infty} d\omega' e^{-(\frac{\omega'-\omega}{\delta\omega})^2} \sigma(\omega') \\
&= c_2 \delta\omega \sum_{i,j} \int_{-\infty}^{\infty} dt e^{(-t/\delta t)^2} e^{-i\omega t} \langle [O_i(t), O_j] \rangle \,,
\end{aligned} \tag{D.21}$$

where $c_1 = \frac{e}{1-e^{-8}}$, $c_2 = c_1\sqrt{\pi}/2$, which we will combine and denote by C, and $\delta t =$

$2/\delta\omega$. The first inequality is because $\sigma(\omega)$ is positive for $\omega > 0$ and $\sigma(-\omega) = -\sigma(\omega)$. The second equality comes from evaluating the integral over ω'. We then use the Lieb-Robinson bound in Ref. [29], which applies for interactions with characteristic exponent $\alpha > d$:

$$\|[O_i(t), O_j(0)]\| \leq Ce^{vt}\left(\frac{1}{r_{ij}^{\alpha}} + e^{-\mu r_{ij}}\right), \tag{D.22}$$

where v, C, μ are positive constants. While this bound was derived in Ref. [29] for 2-body interactions, it also holds for more general k-body interactions and thus for fully general power-law Hamiltonians [see Eq. (D.11)].

We now divide the sum in Eq. (D.21) into two parts corresponding to $r_{ij} > r_*$ and $r_{ij} \leq r_*$ for some parameter r_* we shall choose later. The sum over i, j such that $r_{ij} > r_*$ can then be bounded by first inserting Eq. (D.22) into Eq. (D.21) and evaluating the integration over time. Note that the factor $e^{-t^2/\delta t^2}$ suppresses the contribution from e^{vt} at large t. Therefore, performing the integral yields an upper bound $C(1/r_{ij}^{\alpha} + e^{-\mu r_{ij}})$ for each term corresponding to the pair (i, j), and the sum over $r_{ij} > r_*$ gives:

$$\sum_{i,j:r_{ij}>r_*} C\left(\frac{1}{r_{ij}^{\alpha}} + e^{-\mu r_{ij}}\right) \leq CN\left(\frac{1}{r_*^{\alpha-d}} + e^{-\mu r_*}\right), \tag{D.23}$$

for $\alpha > d$, where the factor of N comes from summing over i and the factor of r^d from summing over j.

On the other hand, for $r_{ij} \leq r_*$, we simply use Eq. (D.15) to bound their contributions. Summing over i, j such that $r_{ij} \leq r_*$, we get a bound $CNr_*^d e^{-\kappa\omega}$, where the factor of N again comes from summing over i and the factor of r_*^d from counting the number of

sites j within a distance r_* from i. Combining with Eq. (D.23) yields an upper bound on the the total heating rate

$$|\sigma([\omega, \omega + \delta\omega])| \leq CNr_*^d \left(e^{-\kappa\omega} + \frac{1}{r_*^\alpha} + r_*^{-d} e^{-\mu r_*} \right). \tag{D.24}$$

Choosing $r_*^\alpha = e^{\kappa\omega}$ and noting that the last term is dominated by the first two when ω is large enough, we find

$$|\sigma([\omega, \omega + \delta\omega])| \leq CNe^{-\frac{\alpha-d}{\alpha}\kappa\omega}, \tag{D.25}$$

which is exponentially small with ω as long as $\alpha > d$.

D.3 The Effective Hamiltonian

In this section, we study the structure of the effective Hamiltonian defined in Eq. (6.116). Specifically, we show that the operators G_q defined in Eq. (6.113) are also power-law Hamiltonians [See also Lemma 7 in the main text for $q < q_{\max}$ and Lemma 26 below for $q \geq q_{\max}$]. In addition, we show that the norm G_q for $q \geq q_{\max}$ is exponentially small as a function of q and ω_* [Lemma 26], implying that the norm of the residual drive V' is also exponentially small.

D.3.1 Structure of G_q for $q < q_{\max}$

First, we prove the statement of Lemma 7 that the operators G_q are also power-law Hamiltonians for all $q < q_{\max}$.

Proof. We proceed by induction and assume that Lemma 7 holds for all q up to $q = q_0 - 1$ for some $q_0 \geq 1$. We now prove that it also holds for $q = q_0$. We consider the first term in the definition of G_{q_0} [Eq. (6.113)]:

$$G_{q_0,1} = \sum_{k=1}^{q_0} \frac{(-1)^k}{k!} \sum_{\substack{1 \leq i_1,\dots,i_k \leq q_0 \\ i_1+\cdots+i_k=q_0}} \mathrm{ad}_{\Omega_{i_1}} \dots \mathrm{ad}_{\Omega_{i_k}} H(t). \tag{D.26}$$

Using Lemma 7 (note that it applies to all $i \leq q_0$) and Lemma 27 in Appendix D.5, we have

$$\begin{aligned}
G_{q_0,1} &\in \sum_{k=1}^{q_0} \frac{1}{k!} \sum_{\substack{1 \leq i_1,\dots,i_k \leq q_0 \\ i_1+\cdots+i_k=q_0}} T^{q_0} c^{q_0} \lambda^{q_0-k} \prod_{j=1}^{k} (i_j - 1)! q_0^k c^{-k} \lambda^k \mathbb{H}_\alpha^{(q_0+1)} \\
&= T^{q_0} c^{q_0} \lambda^{q_0} \sum_{k=1}^{q_0} \frac{q_0^k c^{-k}}{k!} \sum_{\substack{1 \leq i_1,\dots,i_k \leq q_0 \\ i_1+\cdots+i_k=q_0}} \prod_{j=1}^{k} (i_j - 1)! \mathbb{H}_\alpha^{(q_0+1)} \\
&\subseteq T^{q_0} c^{q_0} \lambda^{q_0} \sum_{k=1}^{q_0} \frac{q_0^k c^{-k}}{k!} (q_0 - k)! 2^k \mathbb{H}_\alpha^{(q_0+1)} \\
&\subseteq T^{q_0} c^{q_0} \lambda^{q_0} q_0! \underbrace{\sum_{k=1}^{q_0} \frac{2^k q_0^k c^{-k} (q_0 - k)!}{q_0! k!}}_{\leq c_1} \mathbb{H}_\alpha^{(q_0+1)} \\
&\subseteq c_1 T^{q_0} c^{q_0} \lambda^{q_0} q_0! \mathbb{H}_\alpha^{(q_0+1)},
\end{aligned} \tag{D.27}$$

where c_1 is a constant which exists because the sum over k converges [See Lemma 29 in Appendix D.5].

To get the first equation, we use Lemma 27, with $k_{\max}$ upper bounded by q_0 every time. We have also used the second part of Lemma 28 in the Appendix to bound the sum over $i_1, \dots, i_k$.

Next, we consider the second term in the definition of G_{q_0}:

$$G_{q_0,2}{=}i\sum_{k=1}^{q}\frac{(-1)^{k+1}}{(k{+}1)!}\sum_{\substack{1\le i_1,\dots,i_k,m\le q+1\\ i_1+\cdots+i_k+m=q+1}}\mathrm{ad}_{\Omega_{i_1}}\dots\mathrm{ad}_{\Omega_{i_k}}\partial_t\Omega_m(t). \tag{D.28}$$

Again, we use Lemma 7 and Lemma 27 to show that

$$G_{q_0,2}\in\sum_{k=1}^{q_0}\frac{q_0^k}{(k+1)!}\sum_{\substack{1\le i_1,\dots,i_k,m\le q_0+1\\ i_1+\cdots+i_k+m=q_0+1}}T^{q_0}c^{q_0-k-1}\lambda^{q_0-k-1}\prod_{j=1}^{k}(i_j-1)!(m-1)!\lambda^k\mathbb{H}_\alpha^{(q_0+1)} \tag{D.29}$$

$$=T^{q_0}c^{q_0}\lambda^{q_0}\sum_{k=1}^{q_0}\frac{q_0^kc^{-k}}{(k+1)!}\underbrace{\sum_{\substack{1\le i_1,\dots,i_k,m\le q_0+1\\ i_1+\cdots+i_k+m=q_0+1}}}_{\le 2^{k+1}}\underbrace{\prod_{j=1}^{k}(i_j-1)!(m-1)!}_{\le(q_0+1-(k+1))!=(q_0-k)!}\mathbb{H}_\alpha^{(q_0+1)} \tag{D.30}$$

$$\subseteq T^{q_0}c^{q_0}\lambda^{q_0}2\sum_{k=1}^{q_0}\frac{2^kq_0^kc^{-k}}{(k+1)}(q_0-k)!\mathbb{H}_\alpha^{(q_0+1)} \tag{D.31}$$

$$\subseteq 2T^{q_0}c^{q_0}\lambda^{q_0}q_0!\underbrace{\sum_{k=1}^{q_0}\frac{2^kq_0^kc^{-k}}{k!}\frac{(q_0-k)!}{q_0!}}_{\le c_1}\mathbb{H}_\alpha^{(q_0+1)}\subseteq 2c_1T^{q_0}c^{q_0}\lambda^{q_0}q_0!\mathbb{H}_\alpha^{(q_0+1)}, \tag{D.32}$$

where we have used Lemma 28 in Appendix D.5 to bound the sums over $i_1,\dots,i_k,m$.

Combining Eq. (D.27) and Eq. (D.32), we have

$$G_{q_0}\in 3c_1T^{q_0}c^{q_0}q_0!\lambda^{q_0}\mathbb{H}_\alpha^{(q_0+1)}. \tag{D.33}$$

Note that c_1 can be made arbitrarily small by choosing a larger value for c. Therefore, with c large enough so that $3c_1<1$, we have that Lemma 7 holds for $q=q_0$. □

D.3.2 Structure of G_q for $q \geq q_{\max}$

We now prove Eq. (6.124), which is a similar result to Lemma 7, but for $q \geq q_{\max} = \omega_*$.

Lemma 26. *For all $q \geq q_{max} = \omega_*$, $G_q \in Ce^{-\kappa' q}\mathbb{H}_\alpha$, where C and κ' are constants.*

Proof. Let us first look at the first term in Eq. (6.113):

$$G_{q,1} = \sum_{k=1}^{q} \frac{(-1)^k}{k!} \sum_{\substack{1 \leq i_1,\dots,i_k \leq \omega_* \\ i_1+\cdots+i_k=q}} \mathrm{ad}_{\Omega_{i_1}} \dots \mathrm{ad}_{\Omega_{i_k}} H(t). \tag{D.34}$$

We also recall from Lemma 7 that for all $q \leq \omega_*$,

$$\Omega_q \in T^q(q-1)!c^{q-1}\lambda^{q-1}\mathbb{H}_\alpha^{(q)} = \frac{1}{\lambda c q} T^q q! c^q \lambda^q \mathbb{H}_\alpha. \tag{D.35}$$

For all $q \leq \omega_*$, we have

$$T^q q! c^q \lambda^q \leq (Tc\lambda q)^q \leq (Tc\lambda\omega_*)^q \leq e^{-\kappa q}, \tag{D.36}$$

where we have used $\omega_* = e^{-\kappa}/(Tc\lambda)$. Therefore, for all $q \leq \omega_*$, we have

$$\Omega_q \in \frac{1}{\lambda c q} T^q q! c^q \lambda^q \mathbb{H}_\alpha \in \frac{1}{\lambda c q} e^{-\kappa q} \mathbb{H}_\alpha. \tag{D.37}$$

Note also that $H(t) \in \mathbb{H}_\alpha$. Therefore, using Lemma 27, we have

$$\begin{aligned} \mathrm{ad}_{\Omega_{i_1}} \dots \mathrm{ad}_{\Omega_{i_k}} H(t) &\in \frac{1}{i_1 \dots i_k} \frac{q^k}{c^k} e^{-\kappa q} \lambda^{-k} \lambda^k \mathbb{H}_\alpha \\ &= \frac{1}{i_1 \dots i_k} \frac{q^k}{c^k} e^{-\kappa q} \mathbb{H}_\alpha. \end{aligned} \tag{D.38}$$

Thus, we get for all q:

$$\begin{aligned} G_{q,1} &\in \underbrace{\left(\sum_{k=1}^{q} \frac{q^k}{c^k k!} \sum_{\substack{1 \le i_1, \dots, i_k \le \omega_* \\ i_1 + \dots + i_k = q}} \frac{1}{i_1 \dots i_k} \right)}_{\le e^{q/c} 2^q} e^{-\kappa q} \mathbb{H}_\alpha \\ &\subseteq e^{-(\kappa - \ln 2 - 1/c) q} \mathbb{H}_\alpha. \end{aligned} \tag{D.39}$$

Note that the $i_j \le \omega_*$ as we only define Ω up to ω_*. Further, the factor of 2^q comes from upper-bounding $\frac{1}{i_1 \dots i_k}$ with 1 and the number of terms with 2^q. Next, we consider the second term in the definition of G_q:

$$G_{q,2} = i \sum_{k=1}^{q} \frac{(-1)^{k+1}}{(k+1)!} \sum_{\substack{1 \le i_1, \dots, i_k, m \le q+1 \\ i_1 + \dots + i_k + m = q+1}} \mathrm{ad}_{\Omega_{i_1}} \dots \mathrm{ad}_{\Omega_{i_k}} \partial_t \Omega_m(t). \tag{D.40}$$

Note that

$$\partial_t \Omega_m(t) \in T^{m-1} (m-1)! c^{m-1} \lambda^{m-1} \mathbb{H}_\alpha \subseteq e^{-\kappa(m-1)} \mathbb{H}_\alpha. \tag{D.41}$$

Thus, we have

$$G_{q,2} \in \left(\sum_{k=1}^{q} \frac{q^k}{c^k (k+1)!} \sum_{\substack{1 \le i_1, \ldots, i_k, m \le q+1 \\ i_1 + \cdots + i_k + m = q+1}} e^{-\kappa q} \right) \mathbb{H}_\alpha$$

$$\subseteq 2e^{-(\kappa - \ln 2 - 1/c)q} \mathbb{H}_\alpha. \tag{D.42}$$

Combining Eq. (D.39) and Eq. (D.42), we arrive at Lemma 26 with $\kappa' = \kappa - \ln 2 - 1/c$, which can be made to be positive by choosing $\kappa > \ln 2 + 1/c$. It suffices, however, to choose $\kappa > \ln 2$, since making c large enough sends $1/c$ to zero. Equation (6.125) also follows. □

D.4 Using the Lieb-Robinson Bounds for Evolutions of Local Observables

In this section, we use the Lieb-Robinson bounds to bound the norm of δ in Eq. (6.126). In the main text, we argue that $\|\delta(t)\|$ would be small up to time $t_* \propto \omega_*$ if the light cone induced by the Lieb-Robinson bound is logarithmic, and $t_* \propto e^{\kappa' \omega_*}$ if the light cone is algebraic. We provide below the mathematical details to supplement the argument.

Recall that $V'(t) \in Ce^{-\kappa' \omega_*} \mathbb{H}_\alpha, \bar{H}' \in \gamma \mathbb{H}_\alpha$ (γ is a constant that depends only on κ, α) and that we defined the normalized $V'' = C^{-1} e^{\kappa' \omega_*} V, \bar{H}'' = \gamma^{-1} \bar{H}'$ such that:

$$\|\delta\| \le Ce^{-\kappa' \omega_*} \int_0^t ds \| \left[V''(s), e^{is\gamma \bar{H}''} O e^{-is\gamma \bar{H}''} \right] \|. \tag{D.43}$$

We now use a Lieb-Robinson bound for power-law interactions to bound the commutator.

The idea is that for a finite time s, the operator O mostly spreads within a light cone, and only the terms of $V''(s)$ within the light cone significantly contribute to the commutator.

In contrast to the finite-range interacting Hamiltonians, a tight Lieb-Robinson bound has yet to be proven for power-law Hamiltonians with finite $\alpha > d$. In the following sections, we consider the effect of using different Lieb-Robinson bounds, namely the bounds in Gong et al. [29], Else et al. [141], Tran et al. [21]. The case of a hypothetical bound, which would be tight if it were proven, is treated in the main text.

D.4.1 Using Gong et al.'s Bound

First, we consider a generalization of the bound in Gong et al. [29] [See also Eq. (D.12)]. The bound holds for $\alpha > d$, has a logarithmic light cone $t \gtrsim \log r$, and is extended to many-body interactions. To bound the commutator norm in Eq. (7.5), recall that we write $V''(s) = \sum_{r=0}^{\infty} V_r''(s)$, where $V_r''(s) \equiv \sum_{X:\mathrm{dist}(()X,O)\in[r,r+1)} h_X$ denotes the terms of $V''(s)$ supported on subsets exactly a distance between r and $r+1$ away from O. furthermore, since $V''(s)$ is a power-law Hamiltonian, it follows that $\|V_r''(t)\| \le Cr^{d-1}$.

From Eq. (D.12), the light cone of the bound is $r_*(s) = e^{vs/\alpha}$. We further divide $V_r''(s)$ into those with $r \le r_*(s)$ and $r > r_*(s)$. In the former case, we simply bound:

$$\begin{aligned}&\sum_{r\le r_*(s)} \left\| \left[V_r''(s), e^{is\gamma \bar{H}''} O e^{-is\gamma \bar{H}''} \right] \right\| \\ &\le 2 \sum_{r\le r_*(s)} \|V_r''(s)\|\|O\| \le C r_*(s)^d \le C e^{dvs/\alpha}. \end{aligned} \tag{D.44}$$

For the latter case, we use Eq. (D.12) to bound the commutator norm:

$$\sum_{r>r_*(s)} \left\| \left[V_r''(s), e^{is\gamma \bar{H}''} O e^{-is\gamma \bar{H}''} \right] \right\|$$

$$\leq C \sum_{r>r_*(s)} \|V_r''(s)\| \|O\| \left(\frac{e^{vs}}{r^{\alpha-d}} + e^{vs-\mu r} \right) \tag{D.45}$$

$$\leq C \sum_{r>r_*(s)} \left(\frac{e^{vs}}{r^{\alpha-2d+1}} + r^{d-1} e^{vs-\mu r} \right) \tag{D.46}$$

$$\leq C \left(\frac{e^{vs}}{r_*(s)^{\alpha-2d}} + r_*(s)^{d-1} e^{vs-\mu r_*(s)} \right) \tag{D.47}$$

$$\leq C \left(e^{2dvs/\alpha} + e^{vs\frac{d-1}{\alpha}} e^{vs-\mu e^{vs/\alpha}} \right) \tag{D.48}$$

$$\leq C e^{2dvs/\alpha}, \tag{D.49}$$

where we use the same C to denote different constants that may depend on μ, α. Note that while the bound in Eq. (D.12) is valid for $\alpha > d$, the sum over r converges only when $\alpha > 2d$.

Plugging Eq. (D.44) and Eq. (D.49) into Eq. (7.5) and integrating over s, we have

$$\|\delta\| \leq C e^{-\kappa' \omega_*} e^{2dvt/\alpha}, \tag{D.50}$$

which is the result presented in Sec. 6.9.3. Again, δ is only small for up to time $t_* \propto \omega_* \propto 1/T$, which is expected because the region inside the light cone implied by this bound expands exponentially fast with time.

D.4.2 Using Else et al.'s Bound

Instead of using Gong et al. 's bound, we now use the bound in Else et al. [141], which already holds for many-body interactions. The bound states that when $|X| = 1$,

$$\|[A(t),B]\| \leq C\|A\|\|B\| \left\{ \exp\left(vt - r^{1-\sigma}\right) + \frac{(vt)^{1+d/(1-\sigma)}}{r^{\sigma(\alpha-d)}} \right\}, \tag{D.51}$$

where $1 > \sigma > (d+1)/(\alpha - d + 1)$ is a constant that we can choose. Since our aim is to prove an exponential heating time for α as small as possible, we need the algebraic tail exponent $\sigma(\alpha - d)$ to be as large as possible. So we will assume that we pick some σ very close to 1.

First, let us look at the light cone generated by Eq. (D.51). The first term of the bound gives a light cone $t \gtrsim r^{1-\sigma}$, while the second term gives $t \gtrsim r^{(1-\sigma)\frac{\sigma(\alpha-d)}{d+1-\sigma}}$. Since we are choosing σ close to 1, $\frac{\sigma(\alpha-d)}{d+1-\sigma}$ will be larger than 1 when $\alpha > 2d$. The former light cone, i.e. $t \gtrsim r^{1-\sigma}$, is therefore looser and thus dominates the latter. In the rest of the calculation, we take $r_*(t) = t^{1/(1-\sigma)}$ to be the light cone boundary.

Similar to Eq. (D.44), we get an upper bound for the terms inside the light cone:

$$\begin{aligned} \sum_{r \leq r_*(s)} \| \left[V_r''(s), e^{is\gamma \bar{H}''} O e^{-is\gamma \bar{H}''} \right] \| & \\ \leq 2 \sum_{r \leq r_*(s)} \|V_r''(s)\| \|O\| & \\ \leq C r_*(s)^d \leq C s^{d/(1-\sigma)}. & \end{aligned} \tag{D.52}$$

For the terms outside the light cone, we use Eq. (D.51):

$$\begin{aligned}
&\sum_{r>r_*(s)} \left\| \left[V_r''(s), e^{is\gamma \bar{H}''} O e^{-is\gamma \bar{H}''} \right] \right\| \\
&\leq \sum_{r>r_*(s)} \|V_r''(s)\| \|O\| \left(e^{vs - r^{1-\sigma}} + \frac{(vs)^{1+d/(1-\sigma)}}{r^{\sigma(\alpha-d)}} \right) \\
&\leq C \sum_{r>r_*(s)} \left(r^{d-1} e^{vs - r^{1-\sigma}} + \frac{(vs)^{1+d/(1-\sigma)}}{r^{\sigma(\alpha-d)-d+1}} \right) \\
&\leq C \left(\frac{1}{d} \xi \left(\frac{d}{1-\sigma} \right) e^{vs} r_*^d e^{-r_*^{1-\sigma}} + \frac{(vs)^{1+d/(1-\sigma)}}{r_*^{\sigma(\alpha-d)-d}} \right) \\
&\leq C \left(\xi \left(\frac{d}{1-\sigma} \right) s^{d/(1-\sigma)} + \frac{(vs)^{1+d/(1-\sigma)}}{s^{\frac{\sigma(\alpha-d)-d}{1-\sigma}}} \right) \\
&\leq C\xi \left(\frac{d}{1-\sigma} \right) s^{\frac{d}{1-\sigma}},
\end{aligned} \tag{D.53}$$

where $\xi(x) \equiv \frac{1}{x} 2^x \Gamma(x)$, Γ is the Gamma function, and we again absorb all constants that may depend on d alone into the constant C. We drop the second term in the second to last inequality because for σ arbitrarily close to 1 and $\alpha > 2d$ (see below), the second term may be upper-bounded by the first. To estimate the sum over r, we have used Lemma 30 in Appendix D.5.2. Plugging Eq. (D.52) and Eq. (D.53) into Eq. (7.5) and integrating over time, we get

$$\|\delta\| \leq C e^{-\kappa' \omega_*} \xi \left(\frac{d}{1-\sigma} \right) t^{\frac{d}{1-\sigma}+1}. \tag{D.54}$$

Thus, the difference is small up to an exponentially long time $t_* \propto e^{\kappa' \omega_* \frac{1-\sigma}{d+1-\sigma}}$. The sum over r converges if $\sigma(\alpha - d) > d$, or equivalently $\alpha > d\left(1 + \frac{1}{\sigma}\right)$. Since σ can be chosen arbitrarily close to 1, this condition is effectively equivalent to $\alpha > 2d$.

One should be careful, however, in taking the limit σ goes to one since i) the heating time $t_* \propto e^{\kappa' \omega_* \frac{1-\sigma}{d+1-\sigma}}$ is no longer exponential in ω_* and ii) the prefactor $\xi\left(\frac{d}{1-\sigma}\right)$ diverges faster than exponentially in this limit. Nevertheless, the analysis is still valid for fixed values of $\sigma < 1$.

D.4.3 Using Tran et al.'s Bound

In addition to Else et al. [141]'s bound, we can also use the bound in Tran et al. [21] [see also Eq. (6.104) for a generalization to k-body interactions], which also works for $\alpha > 2d$. Compared to the bound in Else et al. , the bound in Tran et al. has a tighter light cone $r_*(s) = s^{(\alpha-d)/(\alpha-2d)}$, but it decays with the distance r as $r^{\alpha-2d}$, slower than the tail $r^{\sigma(\alpha-d)}$ in Else et al. when $\sigma > (\alpha-2d)/(\alpha-d)$.

Similar to before, we further divide $V_r''(s)$ into those with $r \leq r_*(s)$ and $r > r_*(s)$. For the terms inside the light cone, we again bound:

$$\begin{aligned}
&\sum_{r \leq r_*(s)} \| \left[V_r''(s), e^{is\gamma \bar{H}''} O e^{-is\gamma \bar{H}''} \right] \| \\
&\leq 2 \sum_{r \leq r_*(s)} \|V_r''(s)\| \|O\| \\
&\leq C r_*(s)^d \leq C s^{d(\alpha-d)/(\alpha-2d)}.
\end{aligned} \tag{D.55}$$

For the terms outside the light cone, we use Eq. (6.104) with $\phi(X) = 1$:

$$\begin{aligned}
&\sum_{r>r_*(s)} \left\| \left[V_r''(s), e^{is\gamma \bar{H}''} O e^{-is\gamma \bar{H}''} \right] \right\| \\
&\leq \sum_{r>r_*(s)} \|V_r''(s)\| \|O\| \left(\frac{s^{\alpha-d}}{r^{\alpha-2d}} + s r^{d-1} e^{-r/s} \right) \\
&\leq C \sum_{r>r_*(s)} \left(\frac{s^{\alpha-d}}{r^{\alpha-3d+1}} + s r^{2d-2} e^{-\mu r/s} \right) \\
&\leq C \left(\frac{s^{\alpha-d}}{r_*(s)^{\alpha-3d}} + s^2 r_*(s)^{2d-2} e^{-\mu r_*(s)/s} \right) \\
&\leq C \left(s^{\frac{d(\alpha-d)}{\alpha-2d}} + s^2 s^{2(\alpha-d)(d-1)/(\alpha-2d)} e^{-\mu s^{d/(\alpha-2d)}} \right) \\
&\leq C s^{\frac{d(\alpha-d)}{\alpha-2d}},
\end{aligned} \tag{D.56}$$

where we have dropped the second term in the second to last inequality because it is exponentially small in s and can be upper bounded by the first term. Note that we require $\alpha > 3d$ in order for the sum over r to converge.

Plugging Eq. (D.55) and Eq. (D.56) into Eq. (7.5) and integrating over time, we get

$$\|\delta\| \leq C e^{-\kappa' \omega_*} t^{\frac{d(\alpha-d)}{\alpha-2d}+1}. \tag{D.57}$$

Thus, the difference is small up to an exponentially long time $t_* \propto e^{\kappa' \omega_* \frac{\alpha-2d}{\alpha(d+1)-d(d+2)}}$. Compared to using Else et al. 's bound, this analysis works only when $\alpha > 3d$. However, within this regime, the exponent of the heating time using this bound is larger than using Else et al. . This is a manifestation of the trade-off between the tail and the light cone when switching from Else et al. to Tran et al. bound.

D.5 Mathematical Preliminaries

This section contains mathematical details omitted from the previous sections for clarity. In Appendix D.5.1, we discuss the properties of the set of power-law Hamiltonians defined in Definition 1. In Appendix D.5.2, we present some bounds on discrete sums.

D.5.1 Properties of the Set $\mathbb{H}_\alpha$ of Power-Law Hamiltonians

In this section, we explore some properties of $\mathbb{H}_\alpha$ that are useful for proving that the effective Hamiltonian is also power-law [See Appendix D.3].

We recall from the main text that $\mathbb{H}_\alpha$ is the set of power-law Hamiltonians with the exponent α. In addition, $\mathbb{H}_\alpha^{(k)}$ is the subset of $\mathbb{H}_\alpha$ which contains all power-law Hamiltonians whose local support size [see Definition 1] is at most $k+1$. For a real positive constant a, we also denote by $a\mathbb{H}_\alpha$ the set of Hamiltonians H such that $a^{-1}H$ is a power-law Hamiltonian with the exponent α. It is straightforward to prove the following identities:

$$a\mathbb{H}_\alpha + b\mathbb{H}_\alpha \subset (a+b)\mathbb{H}_\alpha, \tag{D.58}$$

$$a\mathbb{H}_\alpha \subset b\mathbb{H}_\alpha \text{ if } a \leq b. \tag{D.59}$$

The following lemma is particularly useful for the adjoint operation:

Lemma 27. *For $\alpha > d$, if $H_1 \in a\mathbb{H}_\alpha^{(k_1)}, H_2 \in b\mathbb{H}_\alpha^{(k_2)}$ for some positive constants a, b, k_1, k_2, then $\mathrm{ad}_{H_1} H_2 \in ab\lambda k_{\max}\mathbb{H}_\alpha^{(k_1+k_2)}$, where λ is a constant to be defined later and $k_{\max} = \max\{k_1, k_2\}$.*

Proof. Write $H_1 = \sum_X a_X, H_2 = \sum_Y b_Y, \mathrm{ad}_{H_1} H_2 = \sum_Z h_Z$, where $h_Z = \mathrm{ad}_{h_X} h_Y$ and

$Z = X \cup Y$. By our definition of power-law Hamiltonians, we have:

$$\sum_{X \ni i,j} \|a_X\| \leq \frac{a}{\text{dist}\,(()\, i,j)^\alpha}, \quad \sum_{Y \ni i,j} \|b_Y\| \leq \frac{b}{\text{dist}\,(()\, i,j)^\alpha}. \tag{D.60}$$

When $\alpha > d$, it is also straightforward to prove that $\sum_{X \ni i} \|a_X\| \leq a\lambda_0$ for all i, where λ_0 is a constant that depends only on α, d.

Note that $h_Z \neq 0$ only if $X \cap Y \neq \varnothing$. We seek to bound $\sum_{Z \ni i,j} \|h_Z\|$ which sums over $Z = X \cup Y \ni i,j$ and $X \cap Y \neq \varnothing$. We discuss some useful notations. We will occasionally rewrite or label summations with restrictions using the indicator function $\xi(A)$ where $\xi(A) = 1, 0$ if A is true, false respectively. There are 9 mutually exclusive cases [Table D.1], satisfying $i,j \in X \cup Y$ depending on whether i,j are in X, Y, or both.

	$\in X$	$\notin X$	$\in Y$	$\notin Y$
ξ_1	i,j	$-$	i,j	$-$
ξ_2	i,j	$-$	i	j
ξ_3	i,j	$-$	j	i
ξ_4	i,j	$-$	$-$	i,j
ξ_5	i	j	i,j	$-$
ξ_6	i	j	j	i
ξ_7	j	i	i,j	$-$
ξ_8	j	i	i	j
ξ_9	$-$	i,j	i,j	$-$

Table D.1: Mutually exclusive indicator functions for Lemma 27. For example, $\xi_1 = 1$ if all of the conditions in the first row, i.e. $i,j \in X$ and $i,j \notin Y$, hold and $\xi_1 = 0$ otherwise.

Thus, the indicator function $\xi(X \cup Y \ni i,j)$ may be written as a sum of indicator functions of mutually exclusive events listed in the table: $\xi(X \cup Y \ni i,j) = \sum_{n=1}^{9} \xi_n$.

The overall sum that we want to bound can be written as a sum over the nine cases:

$$\begin{aligned}
&\sum_{Z \ni i,j} \|h_Z\| = \sum_{X \cup Y \ni i,j} \|[a_X, b_Y]\| \\
&\leq 2 \sum_X \sum_Y \|a_X\| \|b_Y\| \xi(X \cap Y \neq \varnothing) \xi(X \cup Y \ni i,j) \\
&= 2 \sum_{n=1}^{9} \sum_X \sum_Y \|a_X\| \|b_Y\| \xi(X \cap Y \neq \varnothing) \xi_n,
\end{aligned} \tag{D.61}$$

and we will bound each of the nine cases individually. We will often eliminate the condition that $X \cap Y \neq \varnothing$, which can only make the sum larger, and introduce an inequality by summing over all sets X or Y. To illustrate our technique, consider first the contribution from ξ_5:

$$\begin{aligned}
&2 \sum_X \sum_Y \|a_X\| \|b_Y\| \xi(X \cap Y \neq \varnothing) \xi_5 \\
&\leq 2 \sum_{X \ni i} \sum_{Y \ni i,j} \|a_X\| \|b_Y\| \\
&\leq 2 \sum_{X \ni i} \|a_X\| \frac{b}{\text{dist}\,(() \, i,j)^\alpha} \leq \frac{2\lambda_0 ab}{\text{dist}\,(() \, i,j)^\alpha},
\end{aligned} \tag{D.62}$$

where the first inequality comes from ignoring $j \notin X$ and the second comes from H_2 being a power-law Hamiltonian.

The bound on the term corresponding to ξ_7 follows analogously since we simply switch i, j. Similarly, the terms corresponding to ξ_2, ξ_3 switch only the roles of X, Y

compared to ξ_5, ξ_7. Meanwhile, analyzing the term corresponding to ξ_1 yields:

$$\begin{aligned} & 2\sum_X \sum_Y \|a_X\| \|b_Y\| \xi_1 \xi(X \cap Y \neq \varnothing) \\ & = 2 \sum_{X \ni i,j} \sum_{Y \ni i,j} \|a_X\| \|b_Y\| \\ & \leq \frac{2ab}{\text{dist}\,(()\, i,j)^{2\alpha}} \leq \frac{2ab}{\text{dist}\,(()\, i,j)^{\alpha}}, \end{aligned} \tag{D.63}$$

where we take into account $\text{dist}\,(()\, i,j) \geq 1$ for all d.

Upper bounding the term corresponding to ξ_6 is a bit trickier. Since $X \cap Y \neq \varnothing$, there exists a site $\ell \neq i,j$ such that $\ell \in X \cap Y$. Rewriting the term corresponding to ξ_6 as a sum over ℓ, we have:

$$\begin{aligned} & \sum_{\substack{X \ni i \\ X \not\ni j}} \sum_{\substack{Y \ni j \\ Y \not\ni i}} \|a_X\| \|b_Y\| \xi(X \cap Y \neq \varnothing) \\ & \leq 2 \sum_{\ell \neq i,j} \sum_{X \ni i,\ell} \sum_{Y \ni j,\ell} 2\|a_X\| \|b_Y\| \\ & \leq 2 \sum_{\ell \neq i,j} \frac{a}{\text{dist}\,(()\, i,\ell)^{\alpha}} \frac{b}{\text{dist}\,(()\, \ell,j)^{\alpha}} \leq \frac{2\lambda_1 ab}{\text{dist}\,(()\, i,j)^{\alpha}}, \end{aligned} \tag{D.64}$$

where the last inequality comes from the reproducibility condition [18], applicable when $\alpha > d$, and λ_1 is a constant that depends only on d, α. The term corresponding to ξ_8 contributes the same as ξ_6, as it only switches the roles of i, j.

Finally, we bound the terms corresponding to ξ_4, ξ_9. For ξ_4, we are trying to bound

the sum:

$$\sum_{X \ni i,j} \sum_{Y \not\ni i,j} 2\|a_X\|\|b_Y\|\xi(X \cap Y \neq \varnothing). \tag{D.65}$$

The non-empty intersection means that for there to be a nonzero contribution, $\exists \ell \neq i, j$ such that $\ell \in X, Y$. Further note that by assumption the maximum extent of X is $k_1 + 1$ and therefore there are at most $k_1 - 1$ sites distinct from i, j where Y can intersect with X. We bound this as follows:

$$\begin{aligned}
&2 \sum_{X \ni i,j} \sum_{Y \not\ni i,j} \|a_X\|\|b_Y\|\xi(X \cap Y \neq \varnothing) \\
&\leq 2 \sum_{X \ni i,j} \sum_{\substack{\ell \in X \\ \ell \neq i,j}} \sum_{Y \ni \ell} \|a_X\|\|b_Y\| \\
&\leq 2 \sum_{X \ni i,j} \|a_X\| \sum_{\substack{\ell \in X \\ \ell \neq i,j}} \lambda_0 b \leq \frac{2\lambda_0(k_1 - 1)ab}{\text{dist}\,(()\, i, j)^\alpha}.
\end{aligned} \tag{D.66}$$

We bound the term corresponding to ξ_9 similarly by switching the role of X, Y. Collecting everything, we have the lemma with $\lambda = 2(6\lambda_0 + 2\lambda_1 + 1)$. □

D.5.2 Bounds on Discrete Sums

In this section, we provide bounds on some discrete sums used in the main text.

Lemma 28. *For all $1 \leq k \leq q$, we have the following inequalities:*

$$\sum_{\substack{1\leq i_1,\ldots,i_k\leq q\\ i_1+\cdots+i_k=q}} \prod_{j=1}^{k} i_j! \leq \frac{q!}{(k-1)!}, \tag{D.67}$$

$$\sum_{\substack{0\leq i_1,\ldots,i_k\leq q\\ i_1+\cdots+i_k=q}} \prod_{j=1}^{k} i_j! \leq 2^k q!. \tag{D.68}$$

Proof. We first bound

$$\sum_{\substack{1\leq i_1,\ldots,i_k\leq q\\ i_1+\cdots+i_k=q}} \prod_{j=1}^{k} i_j! \leq \binom{q-1}{k-1} \max_{\substack{1\leq i_1,\ldots,i_k\leq q\\ i_1+\cdots+i_k=q}} \prod_{j=1}^{k} i_j!. \tag{D.69}$$

For positive integers $a \geq b$, we have $(a+b-1)! = a!(a+b-1)\cdots(a+1) \geq a!b!$ with equality if either $a, b = 1$. This implies that the maximal product occurs for some $i_j = q-k+1$ and $i_{k\neq j} = 1$ (we omit the simple proof by induction), yielding

$$\begin{aligned}
&\binom{q-1}{k-1} \max_{\substack{1\leq i_1,\ldots,i_k\leq q\\ i_1+\cdots+i_k=q}} \prod_{j=1}^{k} i_j! \\
&\leq \frac{(q-1)!}{(k-1)!(q-k)!}(q-k+1)! \\
&\leq \frac{(q-1)!}{(k-1)!}(q-k+1) \leq \frac{q!}{(k-1)!},
\end{aligned} \tag{D.70}$$

as $k \neq 0$ by the summation restrictions. Eq. (D.68) is essentially the same as Eq. (D.67) with some indices allowed to be 0. For example, if $i_1 = 0$ while the other i are nonzero, it is just Eq. (D.67) with $k \to k-1$. This part of the sum is then crudely upper bounded by $q!$, while summing over all possible choices of zero indices leads to a factor 2^k. □

Corollary 5. *For all* $1 \leq k \leq q$*, we have:*

$$\sum_{\substack{1 \leq i_1, \ldots, i_k \leq q_0 \\ i_1 + \cdots + i_k = q_0}} \prod_{j=1}^{k} (i_j - 1)! \leq 2^k (q_0 - k)!. \tag{D.71}$$

Proof. Define $p_j = i_j - 1$ such that $0 \leq p_j \leq q_0 - 1$ and $p_1 + \cdots + p_k = q_0 - k$. This second condition implies that we may simplify the first condition to $0 \leq p_j \leq q_0 - k$. Therefore:

$$\begin{aligned} &\sum_{\substack{1 \leq i_1, \ldots, i_k \leq q_0 \\ i_1 + \cdots + i_k = q_0}} \prod_{j=1}^{k} (i_j - 1)! \\ &= \sum_{\substack{0 \leq p_1, \ldots, p_k \leq q_0 - k \\ p_1 + \cdots + p_k = q_0 - k}} \prod_{j=1}^{k} p_j! \leq 2^k (q_0 - k)!, \end{aligned} \tag{D.72}$$

where the last inequality is from Eq. (D.68). □

Lemma 29. *For all* $1 \leq k \leq q$*, we have:*

$$\sum_{k=1}^{q_0} \frac{2^k q_0^k c^{-k} (q_0 - k)!}{q_0! k!} \leq \frac{e}{\sqrt{2\pi}} (e^{2e/c} - 1) \tag{D.73}$$

Proof. Using Stirling's approximation, $\sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} \leq n! \leq e n^{n+\frac{1}{2}} e^{-n}$ for $q_0!$ and $(q_0 -$

$k)!$, we can bound:

$$\begin{aligned}
&\sum_{k=1}^{q_0} \frac{2^k q_0^k c^{-k}}{k!} \frac{(q_0-k)!}{q_0!} \\
&\leq \sum_{k=1}^{q_0} \frac{2^k q_0^k c^{-k}}{k!} \frac{e}{\sqrt{2\pi}} \frac{(q_0-k)^{q_0-k}}{q_0^{q_0}} \frac{\sqrt{q_0-k}}{\sqrt{q_0}} \frac{e^{-(q_0-k)}}{e^{-q_0}} \\
&\leq \frac{e}{\sqrt{2\pi}} \sum_{k=1}^{q_0} \frac{2^k e^k c^{-k}}{k!} \underbrace{\frac{(q_0-k)^{q_0-k}}{q_0^{q_0-k}}}_{\leq 1} \underbrace{\frac{\sqrt{q_0-k}}{\sqrt{q_0}}}_{\leq 1} \\
&\leq \frac{e}{\sqrt{2\pi}} \sum_{k=1}^{\infty} \frac{2^k e^k c^{-k}}{k!} = \frac{e}{\sqrt{2\pi}} (e^{2e/c} - 1).
\end{aligned} \tag{D.74}$$

We note that the bound approaches 0 as $c \to \infty$. □

Lemma 30. *For $d \in \mathbb{N}_{>0}, r_* > 1, 0 < \eta < 1$*

$$\sum_{r > r_*} r^{d-1} e^{-r^\eta} \leq \frac{2}{\eta} 2^{d/\eta} \Gamma(d/\eta) r_*^d e^{-r_*^\eta}, \tag{D.75}$$

where Γ is the Gamma function.

Proof. Let $f(r) = r^{d-1} e^{-r^\eta}$. Our strategy is to upper bound $\sum_{r>r_*} f(r)$ by an integral. For $r \in (0, \infty)$, f has a maximum at $r = r_0 = (d-1)^{1/\eta} \eta^{-1/\eta}$. Let $r_0^- = \lfloor r_0 \rfloor$ and $r_0^+ = r_0^- + 1 > r_0$. Then, the function $f(r)$ is increasing for $r \in (r_*, r_0^-)$ and decreasing

for $r \geq r_0^+$. Therefore, we can upper bound:

$$\begin{aligned}\sum_{r>r_*} f(r) &\leq \int_{r_*}^{r_0^-} f(r)dr + \int_{r_0^+}^{\infty} f(r)dr + f(r_0^-) + f(r_0^+) \\ &\leq \int_{r_*}^{r_0^-} f(r)dr + \int_{r_0^+}^{\infty} f(r)dr + 2\int_{r_0^-}^{r_0^+} f(r)dr \\ &\leq 2\int_{r_*}^{\infty} f(r)dr, \end{aligned} \tag{D.76}$$

where we use the fact that $f(r)$ is concave between r_0^- and r_0^+ to bound the first line by the second line. Next, to bound the integral, we make a change of variable to $x = r^\eta$ so that

$$\begin{aligned}2\int_{r_*}^{\infty} f(r)dr &= 2\int_{r_*}^{\infty} r^{d-1}e^{-r^\eta}dr \\ &= \frac{2}{\eta}\int_{x_*}^{\infty} x^{\frac{d-\eta}{\eta}}e^{-x}dx \\ &\leq \frac{2}{\eta}\int_{x_*}^{\infty} x^\beta e^{-x}dx \leq \frac{2}{\eta}2^\beta\beta!x_*^\beta e^{-x_*} \\ &= \frac{2}{\eta}2^\beta\beta!r_*^{\eta\beta}e^{-r_*^\eta} \\ &\leq \frac{2}{\eta}2^{d/\eta}\Gamma(d/\eta)r_*^d e^{-r_*^\eta}, \end{aligned} \tag{D.77}$$

where $x_* = r_*^\eta$, $\beta = \lceil (d-\eta)/\eta \rceil \leq d/\eta$ is an integer, and Γ is the Gamma function. Note that we have also used a bound for the integral

$$\int_{x_*}^{\infty} x^\beta e^{-x}dx \leq 2^\beta\beta!x_*^\beta e^{-x_*}, \tag{D.78}$$

which can be proven inductively on β for all $\beta \geq 0$ and $x_* \geq 2$. Indeed, the inequality is

trivial for $\beta = 0$. Suppose the inequality holds for $\beta - 1$, using integration by parts, we have

$$\begin{aligned}
\int_{x_*}^{\infty} x^{\beta} e^{-x} dx &= x_*^{\beta} e^{-x_*} + \beta \int_{x_*}^{\infty} x^{\beta-1} e^{-x} dx \\
&\leq x_*^{\beta} e^{-x_*} + \beta 2^{\beta-1} (\beta - 1)! x_*^{\beta-1} e^{-x_*} \\
&\leq 2^{\beta-1} \left(\frac{1}{2^{\beta-1}\beta!} + \frac{1}{x_*} \right) \beta! x_*^{\beta} e^{-x_*} \\
&\leq 2^{\beta} \beta! x_*^{\beta} e^{-x_*}, \qquad \text{(D.79)}
\end{aligned}$$

where the terms inside the bracket in the second to last line is always less than or equal to 2 for all $x_* \geq 1$ (corresponding to $r_* > 1$). □

Appendix E: Supplemental Material for Chapter 7

In this Appendix, we provide more mathematical details for the derivations of the error bound in the Chapter 7. Specifically, Appendix E.1 explains how we write the k-th order error δ_k into a commutator. Appendix E.2 provides an upper bound for a sum of different evolutions of δ. Finally, in Appendix E.3, we show how we bound the norm of Δ_k in Eq. (7.16).

E.1 Structure of δ_k

In this section, we present the proof of Lemma 8, which says that we can write δ_k into a sum of a commutator and an operator of higher order. First, we need the following recursive relation between the δ_k operators.

Lemma 31. *For $k \geq 2$, we have the following recursive relation:*

$$\delta_{k+1} = H_1\delta_k + \delta_k H_2 - [H^k, H_2]. \tag{E.1}$$

Proof. We prove the lemma by expanding both $U_{t/r}$ and $U_{t/r}^{(1)}U_{t/r}^{(2)}$ in orders of t/r:

$$U_{t/r}^{(1)}U_{t/r}^{(2)} = e^{-iH_1 t/r}e^{-iH_2 t/r} = \sum_{k=0}^{\infty} \frac{1}{k!} A_k \left(\frac{-it}{r}\right)^k, \tag{E.2}$$

$$U_{t/r} = e^{-iHt/r} = \sum_{k=0}^{\infty} \frac{1}{k!} B_k \left(\frac{-it}{r}\right)^k, \tag{E.3}$$

where

$$A_k \coloneqq \sum_{j=0}^{k} \binom{k}{j} H_1^j H_2^{k-j}, \quad B_k \coloneqq H^k = (H_1 + H_2)^k. \tag{E.4}$$

With these notations, we have the relation $\delta_k = B_k - A_k$. It is also straightforward to verify the recursive relations for A_k and B_k:

$$A_{k+1} = H_1 A_k + A_k H_2, \tag{E.5}$$

$$\begin{aligned} B_{k+1} = H^{k+1} = HB_k &= (H_1 + H_2)(A_k + \delta_k) \\ &= H_1 A_k + H_1 \delta_k + B_k H_2 - [B_k, H_2] \\ &= H_1 A_k + H_1 \delta_k + (A_k + \delta_k) H_2 - [B_k, H_2] \\ &= (H_1 A_k + A_k H_2) + H_1 \delta_k + \delta_k H_2 - \left[H^k, H_2\right] \\ &= A_{k+1} + H_1 \delta_k + \delta_k H_2 - \left[H^k, H_2\right]. \end{aligned} \tag{E.6}$$

By definition, we have

$$\delta_{k+1} = B_{k+1} - A_{k+1} = H_1 \delta_k + \delta_k H_2 - \left[H^k, H_2\right]. \tag{E.7}$$

Therefore, the lemma follows. □

We now construct the operators S_k, V_k in Lemma 8 inductively on k. For $k = 2$, we have $\delta_2 = [H, H_2]$. Thus Lemma 8 is true for $k = 2$ with $S_2 = H_2$ and $V_2 = 0$. Assume that Lemma 8 is true up to k, i.e. there exist S_k, V_k such that $\delta_k = [H, S_k] + V_k$, we shall prove that it is also true for $k + 1$. Using Lemma 31, we have

$$\begin{aligned}\delta_{k+1} &= H_1\delta_k + \delta_k H_2 - [H^k, H_2] \\ &= [H_1, \delta_k] + \delta_k H - [H^k, H_2] \\ &= [H_1, [H, S_k] + V_k] + V_k H + [H, S_k]H - [H^k, H_2]. \end{aligned} \tag{E.8}$$

We use the following commutator identities:

$$[H, S_k]H = [H, S_k H], \tag{E.9}$$

$$[H^k, H_2] = [H, \sum_{j=0}^{k-1} H^{k-1-j} H_2 H^j]. \tag{E.10}$$

With some trivial manipulations, we can write $\delta_{k+1} = [H, S_{k+1}] + V_{k+1}$, where

$$S_{k+1} = S_k H - \sum_{j=0}^{k-1} H^{k-1-j} H_2 H^j, \tag{E.11}$$

$$V_{k+1} = [H_1, [H, S_k]] + H_1 V_k + V_k H_2. \tag{E.12}$$

Finally, we show that the operators S_k, V_k constructed using the above recursive relations satisfy the norm bounds in Eqs. (7.10)–(7.12). We need the following lemma about the structure of S_k, V_k.

Lemma 32. *For integer $k \geq 2$, the operators S_k, V_k constructed from Eqs.* (E.11) *and* (E.12) *can be written as*

$$V_k = \sum_{i=1}^{n_k} v_{k,i}, \quad n_k \leq \xi e^{k-2} n^{k-2}, \tag{E.13}$$

$$S_k = \sum_{i=1}^{m_k} s_{k,i}, \quad m_k \leq \frac{k(k-1)}{2} n^{k-1}, \tag{E.14}$$

where ξ is a constant, $v_{k,i}, s_{k,i}$ are operators supported on at most $2(k-1)$ sites and

$$\|s_{k,i}\| \leq 1, \quad \|v_{k,i}\| \leq 1, \tag{E.15}$$

for all i.

Proof. Denote by $\mathrm{supp}(X)$ the support size of an operator X, i.e. the number of sites X acts nontrivially on. We say that the number of terms in V_k is x if there exists a decomposition $V_k = \sum_{j=1}^{x} v_j$ such that $\|v_j\| \leq 1$ for all j. For $k = 2$, the lemma is true by definition. Assume that the lemma is true up to some $k \geq 2$, we shall prove that it holds for $k+1$.

First, we argue for the bounds on the number of terms m_{k+1}, n_{k+1} in S_{k+1}, V_{k+1} respectively. Since there are m_k terms in S_k, using Eq. (E.11), it is straightforward to bound m_{k+1}—the number of terms in S_{k+1}:

$$m_{k+1} \leq m_k n + k n^k \leq \frac{k(k-1)}{2} n^k + k n^k = \frac{k(k+1)}{2} n^k. \tag{E.16}$$

To bound n_{k+1}, the number of terms in V_{k+1}, we use Eq. (E.12) and note that $s_{k,i}$ can non-commute with at most $2\text{supp}(s_{k,i}) = 4(k-1)$ terms from H. Therefore, the number of terms in $[H, S_k]$ is at most $4(k-1)m_k$. Each of these terms has its support size increased by at most one (to $2k-1$) compared to the terms of S_k. Repeating the argument for $[H_1, [H, S_k]]$, the number of terms in V_{k+1} can be bounded as follow:

$$n_{k+1} \leq 2(2k-1)4(k-1)m_k + nn_k \tag{E.17}$$

$$\leq 8k^4 n^{k-1} + \xi e^{k-2} n^{k-1} \tag{E.18}$$

$$< 2\xi e^{k-2} n^{k-1} < \xi e^{k-1} n^{k-1}, \tag{E.19}$$

where $\xi = \frac{2048}{e^2(e-1)}$ and we have used the fact that $8k^4 + \xi e^{k-2} < \xi e^{k-1}$ for all $k \geq 2$. Therefore, the number of terms n_{k+1}, m_{k+1} are bounded according to Eqs. (E.13) and (E.14).

It is also apparent from this construction that each iteration in Eqs. (E.11) and (E.12) increases the support size of the constituent terms in S_k, V_k by at most 2. Therefore, Lemma 32 follows. □

With Lemma 32, it is straightforward to show that the norms of $V_k, S_k, [H, S_k]$ are upper bounded by the their number of terms:

$$\|V_k\| \leq n_k = \mathcal{O}\left(e^{k-2} n^{k-2}\right) \tag{E.20}$$

$$\|S_k\| \leq m_k = \mathcal{O}\left(k^2 n^{k-1}\right), \tag{E.21}$$

$$\|[H, S_k]\| \leq 4(k-1)m_k = \mathcal{O}\left(k^3 n^{k-1}\right). \tag{E.22}$$

These bounds complete the proof of Lemma 8.

E.2 Sum of Evolutions of δ

In this section, we present the proof of Lemma 9, which provides an upper bound for the sum of evolution of an operator with different times.

Proof. We denote by $\tau \coloneqq t/r$ and

$$\Sigma_a(X) \coloneqq \sum_{j=0}^{a-1} U_{j\tau} \left[H, X\right] U_{j\tau}^\dagger \tau, \tag{E.23}$$

where X is an arbitrary time-independent operator, a is a positive integer, and $U_t = \exp(-iHt)$ as before.

First, we need to turn the sum $\Sigma_a(X)$ into a sum of several integrals using the following lemma.

Lemma 33. *Define*

$$F[X] \coloneqq -\frac{1}{\tau} \int_0^\tau ds \int_0^s dv U_v \left[H, X\right] U_v^\dagger, \tag{E.24}$$

$$I_t(X) \coloneqq \int_0^t U_s \left[H, X\right] U_s^\dagger ds. \tag{E.25}$$

For all τ such that $n\tau < 1$, where n is the number of sites in the system, we have

$$\Sigma_a(X) = \sum_{k=0}^{\infty} I_{a\tau}(F^{\circ k}[X])), \tag{E.26}$$

where $F^{\circ k}$ the k-th iterate of a function F, i.e. the composition $F^{\circ k}[X] = F[F[\dots F[X]\dots]]$,

with $F^{\circ 0}$ being the identity function.

Proof. To prove the claim, we note that

$$I_{a\tau}(X) = \int_0^{a\tau} U_s [H, X] U_s^\dagger ds = \sum_{j=0}^{a-1} \int_{j\tau}^{(j+1)\tau} U_s [H, X] U_s^\dagger ds \tag{E.27}$$

$$= \sum_{j=0}^{a-1} U_{j\tau} \left(\int_0^{\tau} U_s [H, X] U_s^\dagger ds \right) U_{j\tau}^\dagger. \tag{E.28}$$

Therefore, we have

$$\begin{aligned}
\Sigma_a(X) - I_{a\tau}(X) &= \sum_{j=0}^{a-1} U_{j\tau} \left([H, X]\, \tau - \int_0^{\tau} U_s [H, X] U_s^\dagger ds \right) U_{j\tau}^\dagger \\
&= \sum_{j=0}^{a-1} U_{j\tau} \int_0^{\tau} ds \left([H, X] - U_s [H, X] U_s^\dagger \right) U_{j\tau}^\dagger \\
&= \sum_{j=0}^{a-1} U_{j\tau} \int_0^{\tau} ds \int_s^0 dv U_v [H, [H, X]] U_v^\dagger U_{j\tau}^\dagger \\
&= \sum_{j=0}^{a-1} U_{j\tau} \left[H, \frac{1}{\tau} \int_0^{\tau} ds \int_s^0 dv U_v [H, X] U_v^\dagger \right] U_{j\tau}^\dagger \tau \\
&= \Sigma_a(F[X])).
\end{aligned} \tag{E.29}$$

To get the second last line, we use the fact that H and U_t commute in order to move the integral inside the commutator. Repeated applications of this recursive relation yields Eq. (E.26). The condition $n\tau < 1$ ensures that the sum in Eq. (E.26) converges (See Lemma 34). □

Lemma 34 below is a consequence of Lemma 33.

Lemma 34. *If X is time-independent and $\mu := \frac{nt}{r} < 1$, $\|\Sigma_a(X)\| \leq \frac{2}{1-\mu}\|X\|$.*

Proof. To prove Lemma 34, we note that

$$\|F[X]\| \leq \tau \|H\| \|X\| \leq \mu \|X\|. \tag{E.30}$$

Therefore, $\|F^{\circ k}[X]\| \leq \mu^k \|X\|$. Note also that for the time-independent X,

$$I_{a\tau}(X) = \int_0^{a\tau} U_s \left[H, X\right] U_s^\dagger ds = U_{a\tau} X U_{a\tau}^\dagger - X, \tag{E.31}$$

and therefore $\|I_{a\tau}(X)\| \leq 2\|X\|$. Using Lemma 33, we have

$$\begin{aligned} \|\Sigma_a(X)\| &\leq \sum_{k=0}^{\infty} \|I_{a\tau}(F^{\circ k}[X])\| \leq 2 \sum_{k=0}^{\infty} \|F^{\circ k}[X]\| \\ &\leq 2\|X\| \sum_{k=0}^{\infty} \mu^k = \frac{2}{1-\mu} \|X\| \\ &= \mathcal{O}\left(\|X\|\right), \end{aligned} \tag{E.32}$$

where we have assumed $\mu = \frac{nt}{r} < 1$ so that the sum converges. Therefore, the lemma follows. □

To prove the Lemma 9, we write $\delta = [H, S] + V$ with S, V bounded by Eq. (7.14). We then use Lemma 34 with $X = S$ and the triangle inequality to get

$$\left\| \sum_{j=0}^{a-1} U_{j\tau} \delta \, U_{j\tau}^\dagger \right\| \leq \left\| \frac{1}{\tau} \Sigma_a(S) \right\| + \left\| \sum_{j=0}^{a-1} U_{j\tau} V \, U_{j\tau}^\dagger \right\| \tag{E.33}$$

$$= \mathcal{O}\left(\frac{1}{\tau} \|S\| \right) + \mathcal{O}\left(a \|V\|\right) \tag{E.34}$$

$$= \mathcal{O}\left(\frac{nt}{r} \right) + \mathcal{O}\left(a \frac{nt^3}{r^3} \right). \tag{E.35}$$

Thus, the lemma follows. □

E.3 Upper Bound on Δ_k

In this section, we show how we bound the norms of Δ_k in Eq. (7.16). For that, we use Lemma 9 together with the bound on $\|\delta\|$ [Eq. (7.13)]:

$$\begin{aligned}
\|\Delta_k\| &= \left\| \sum_{i_1=0}^{r-k} \sum_{i_2=0}^{r-k-i_1} \sum_{i_3=0}^{r-k-i_1-i_2} \cdots \sum_{i_k=0}^{r-k-i_1-i_2-\cdots-i_k} \underbrace{U_{t/r}^{i_1} \delta U_{t/r}^{i_2} \delta U_{t/r}^{i_3} \delta \cdots}_{\delta \text{ appears } k \text{ times}} U_{t/r}^{r-k-i_1-i_2-\cdots-i_k} \right\| \\
&\leq \sum_{i_1=0}^{r-k} \sum_{i_2=0}^{r-k-i_1} \sum_{i_3=0}^{r-k-i_1-i_2} \cdots \|\delta\|^{k-1} \left\| \sum_{i_k=0}^{r-k-i_1-i_2-\cdots-i_k} U_{t/r}^{i_k} \delta U_{t/r}^{-i_k} \right\| \\
&\leq r^{k-1} \|\delta\|^{k-1} \mathcal{O}\left(\frac{nt}{r} + \frac{nt^3}{r^2} \right). \qquad \text{(E.36)}
\end{aligned}$$

Thus, Eq. (7.16) follows.

Appendix F: Supplemental Material for Chapter 8

F.1 Faster Convergence of the Quantum Zeno Effect

Using symmetries to protect quantum simulations has previously been explored in the context of the quantum Zeno effect: undesirable errors from the simulation can be suppressed by constantly measuring the system in an appropriate basis [165, 172, 204]. However, measurements are costly in most available quantum computers and therefore often only performed once at the end in simulations on quantum computers. An alternative approach commonly used in quantum control is to frequently apply fast pulses, or "kicks", to the system during the experiments. In the high frequency limit, these kicks confine the dynamics of the system to the so-called quantum Zeno subspaces defined by the spectral decomposition of the kicks [165, 166, 167, 168, 169, 170, 171], effectively realizing the quantum Zeno effect without measuring the systems.

In this section, we derive a concrete bound on the rate at which the effective Hamiltonian of a frequently kicked system converges to its projection to the Zeno subspaces. This bound exponentially improves a recent result of Burgarth, Facchi, Gramegna, and Pascazio [171]. Interestingly, our proof makes use of a tight analysis of Trotter error [77], suggesting a deep connection between quantum simulation and quantum Zeno effect.

The aim of quantum control is to confine the dynamics of a system evolving under

a Hamiltonian G into the subspaces specified by a set of projectors:

$$\mathcal{P} = \{P_\mu\}. \tag{F.1}$$

One approach is to repeatedly measure the system in the basis corresponding to $\mathcal{P}$ throughout the evolution. These measurements results in the quantum Zeno effect: the dynamics of the system is confined to the subspaces corresponding to the projectors P_μ. Alternative to measuring the system, one could periodically "kick" the system [171] with a unitary

$$U_{\text{kick}} = \sum_\mu e^{-i\phi_\mu} P_\mu, \tag{F.2}$$

where ϕ_μ is chosen such that $\phi_\mu \neq \phi_\nu \mod 2\pi$ for all $\mu \neq \nu$.

Suppose the total evolution time is t and we apply the kick every $\delta t = t/r$ where r is an integer. The dynamics of the system becomes

$$(U_{\text{kick}}^\dagger)^r \left(e^{-iGt/r} U_{\text{kick}}\right)^r, \tag{F.3}$$

where $(U_{\text{kick}}^\dagger)^r$ is added at the end of the sequence to undo the evolution generated by the r applications of U_{kick}. In the limit $r \to \infty$, the dynamics of the system again exhibits the quantum Zeno effect

$$U_{\text{kick}}^{\dagger r} \left(e^{-i\frac{t}{r}G} U_{\text{kick}}\right)^r \to e^{-itG_{\text{Zeno}}}, \tag{F.4}$$

where

$$G_{\text{Zeno}} \equiv \sum_{\mu=1}^{m} P_\mu G P_\mu, \tag{F.5}$$

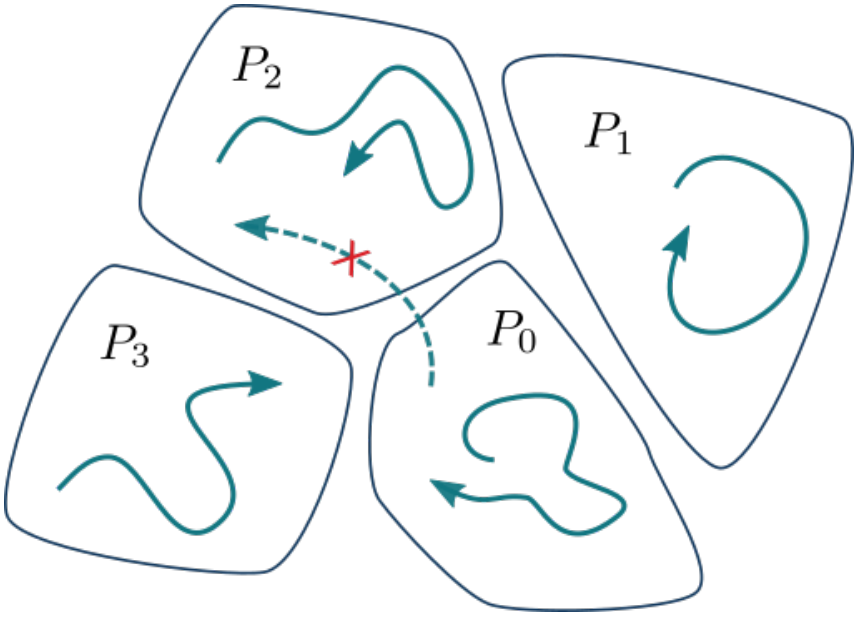

Figure F.1: The frequent kicks confine the dynamics of the system (solid arrows) to the so-called quantum Zeno subspaces, defined by the projectors P_μ in the spectral decomposition of the kicks $U_{\text{kick}} = \sum_\mu e^{-i\phi_\mu} P_\mu$. In particular, the kicks suppress the probability for the system to travel between the subspaces (dashed arrow). By generating the kicks from the symmetries of the system, we can target the simulation error—the sole contributor to possible violations of the symmetries in an ideal simulation—for suppression.

is the projection of G onto the subspaces defined by the spectral decomposition of U_{kick}. In other words, the kicks effectively confine the dynamics of the system to the subspaces defined by the projectors P_μ (See Fig. F.1).

Ref. [171, (A.30)] derived the following bound on the convergence rate with explicit dependence on all parameters of interest

$$\|U_{\text{kick}}^{\dagger r}\left(U_{\text{kick}} e^{-i\frac{t}{r}G}\right)^r - e^{-itG_{\text{Zeno}}}\| \leq \frac{\xi m^2 \|G\| t(1+2e^{m\|G\|t})}{r}, \tag{F.6}$$

where m is the number of projectors and

$$\xi \equiv \max_{\mu\neq\nu} |\sin\left(\frac{\phi_\nu - \phi_\mu}{2}\right)|^{-1} \tag{F.7}$$

is the inverse spectral gap. Unfortunately, this bound has exponential dependence on m, $\|G\|$, and t, which, in particular, suggests that we have to increase the number of kicks r exponentially with the evolution time of the system and therefore may be impractical in many applications. In Theorem 13, we prove a different bound that exponentially improves the bound in Ref. [171] in terms of m, $\|G\|$, and t.

Theorem 13 (Faster convergence of quantum Zeno effect)**.** *Let U_{kick} be the unitary defined in Eq.* (F.2) *with m distinct eigenvalues, inverse spectral gap ξ, and a set of orthogonal projectors $\{P_\mu\}$. Let $G_{Zeno} = \sum_\mu P_\mu G P_\mu$ denote the projection of a Hamiltonian G onto the subspaces defined by $\{P_\mu\}$. We have*

$$\varepsilon_{Zeno} \equiv \|U_{kick}^{\dagger r}\left(U_{kick}e^{-i\frac{t}{r}G}\right)^r - e^{-itG_{Zeno}}\| \leq \frac{2\xi\sqrt{m}\|G\|^2t^2\log r}{r} + \frac{\xi\sqrt{m}\|G\|t}{r} \tag{F.8}$$

$$\leq \frac{3\xi\sqrt{m}\|G\|^2t^2\log r}{r}. \tag{F.9}$$

To prove Theorem 13, we rewrite the evolution as

$$U_{\text{kick}}^{\dagger r}\left(e^{-i\frac{t}{r}G}U_{\text{kick}}\right)^r = e^{-i\frac{t}{r}G_r}e^{-i\frac{t}{r}G_{r-1}}\cdots e^{-i\frac{t}{r}G_1}, \tag{F.10}$$

where we have defined

$$G_k \equiv U_{\text{kick}}^{\dagger k}GU_{\text{kick}}^k. \tag{F.11}$$

Letting $G_{[1,r]} \equiv G_1 + \cdots + G_r$, the first step of our proof is to establish the error bound

$$\|e^{-i\frac{t}{r}G_r}\cdots e^{-i\frac{t}{r}G_1} - e^{-i\frac{t}{r}G_{[1,r]}}\| \leq \frac{2\xi\sqrt{m}\|G\|^2t^2\log r}{r}. \tag{F.12}$$

This is the spectral-norm error of the first-order Trotter formula [77]. However, a naive error analysis in terms of the commutators between G_j (see [77, Proposition 15] for example) gives a bound that does not decrease with r and thus fails to establish the desirable bound. Instead, we seek a better analysis that exploits the spectral information of U_{kick} [171].

The starting point of our analysis is the established von Neumann's ergodic theorem whose proof is included for completeness.

Theorem 14 (Von Neumann's ergodic theorem)**.** *Let U be a unitary operator and $U = \sum_{\mu=1}^{m} e^{-i\phi_\mu} P_\mu$ be its spectral decomposition, with $\phi_1 = 0$ and ϕ_μ distinct. Then,*

$$\|\frac{1}{r}\sum_{k=1}^{r} U^k - P_1\| \leq \frac{\xi_1}{r}, \tag{F.13}$$

where

$$\xi_1 := 2\max_{\nu\neq 1} |e^{-i\phi_\nu} - 1|^{-1} = \max_{\nu\neq 1} |\sin\left(\frac{\phi_\nu}{2}\right)|^{-1}. \tag{F.14}$$

Proof. The bound follows from

$$\begin{aligned}
\|\frac{1}{r}\sum_{k=1}^{r} U^k - P_1\| &= \|\left(\frac{1}{r}\sum_{k=1}^{r} U^k - P_1\right)\sum_{\nu=1}^{m} P_\nu\| \\
&= \|\frac{1}{r}\sum_{\nu=1}^{m}\sum_{k=1}^{r} e^{-ik\phi_\nu} P_\nu - P_1\| \\
&= \|\frac{1}{r}\sum_{\nu\neq 1} e^{-i\phi_\nu}\frac{1-e^{-ir\phi_\nu}}{1-e^{-i\phi_\nu}} P_\nu\| \\
&= \frac{1}{r}\max_{\nu\neq 1}|\frac{1-e^{-ir\phi_\nu}}{1-e^{-i\phi_\nu}}| \leq \frac{\xi_1}{r}.
\end{aligned} \tag{F.15}$$

□

We note that the condition $\phi_1 = 0$ is not restrictive as we can always make $\phi_1 = 0$ by adding a global phase to U_{kick} [167].

Corollary 6. *Let U be a unitary operator and $U = \sum_{\mu=1}^{m} e^{-i\phi_\mu} P_\mu$ be its spectral decomposition. Then, for any operator G,*

$$\|\frac{1}{r}\sum_{k=1}^{r} U^k G U^{-k} - \sum_{\mu=1}^{m} P_\mu G P_\mu\| \leq \frac{\xi\sqrt{m}\|G\|}{r}, \tag{F.16}$$

where

$$\xi := 2\max_{\mu\neq\nu} |e^{-i\phi_\mu} - e^{-i\phi_\nu}|^{-1} = \max_{\mu\neq\nu} |\sin\left(\frac{\phi_\nu - \phi_\mu}{2}\right)|^{-1}. \tag{F.17}$$

Proof. The claimed bound follows from

$$\begin{aligned}
\|\frac{1}{r}\sum_{k=1}^{r} U^k G U^{-k} - \sum_{\mu=1}^{m} P_\mu G P_\mu\| &= \|\left(\frac{1}{r}\sum_{k=1}^{r} U^k G U^{-k} - \sum_{\mu=1}^{m} P_\mu G P_\mu\right)\sum_{\nu=1}^{m} P_\nu\| \\
&= \|\sum_{\nu=1}^{m}\frac{1}{r}\sum_{k=1}^{r}\left(e^{i\phi_\nu}U\right)^k G P_\nu - \sum_{\nu=1}^{m} P_\nu G P_\nu\| \\
&\leq \sqrt{\sum_{\nu=1}^{m}\|\frac{1}{r}\sum_{k=1}^{r}\left(e^{i\phi_\nu}U\right)^k G P_\nu - P_\nu G P_\nu\|^2} \\
&\leq \sqrt{m}\|G\|\max_\nu\|\frac{1}{r}\sum_{k=1}^{r}\left(e^{i\phi_\nu}U\right)^k - P_\nu\| \leq \frac{\xi\sqrt{m}\|G\|}{r},
\end{aligned} \tag{F.18}$$

where the first inequality follows from the bound

$$\begin{aligned}
\|\sum_{\nu=1}^{m} A_\nu P_\nu\| &= \sqrt{\|\left(\sum_{\nu=1}^{m} A_\nu P_\nu\right)\left(\sum_{\mu=1}^{m} A_\mu P_\mu\right)^\dagger\|} \\
&= \sqrt{\|\sum_{\nu=1}^{m} A_\nu P_\nu A_\nu^\dagger\|} \leq \sqrt{\sum_{\nu=1}^{m}\|A_\nu P_\nu A_\nu^\dagger\|} = \sqrt{\sum_{\nu=1}^{m}\|A_\nu P_\nu\|^2}.
\end{aligned} \tag{F.19}$$

□

As aforementioned, a naive analysis of the Trotter error fails to provide the desirable bound for quantum Zeno effect. Instead, we use a recursive approach to estimate the Trotter error Eq. (F.12).

Lemma 35. *Define* $G_{[k_0,k_1]} \equiv \sum_{k=k_0}^{k_1} G_k$ *for* $k_0 \leq k_1$. *For any* $s \geq 1$ *and* δt, *we have*

$$\|\prod_{k=1}^{s} e^{-iG_k\delta t} - e^{-iG_{[1,s]}\delta t}\| \leq 2\xi\sqrt{m}\|G\|^2\delta t^2 s\log_2 s. \tag{F.20}$$

Note that at $s = r$ and $\delta t = t/r$, Lemma 35 reduces to Eq. (F.12). We prove Lemma 35 by induction on s. Suppose Lemma 35 holds for $s = s_1$ and $s = s_2$ such that $|s_2 - s_1| \leq 1$, we shall prove that it holds for $s = s_1 + s_2$. Using the triangle inequality

$$\begin{aligned}
&\|\prod_{k=1}^{s_1+s_2} e^{-iG_k\delta t} - e^{-iG_{[1,s_1+s_2]}\delta t}\| \\
&\leq \|\prod_{k=1}^{s_1} e^{-iG_k\delta t} - e^{-iG_{[1,s_1]}\delta t}\| + \|\prod_{k=s_1+1}^{s_1+s_2} e^{-iG_k\delta t} - e^{-iG_{[s_1+1,s_1+s_2]}\delta t}\| \\
&\quad + \|e^{-iG_{[1,s_1+s_2]}\delta t} - e^{-iG_{[s_1+1,s_1+s_2]}\delta t}e^{-iG_{[1,s_1]}\delta t}\| \qquad \text{(F.21)} \\
&\leq 2\xi\sqrt{m}\|G\|^2\delta t^2(s_1\log_2 s_1 + s_2\log_2 s_2) + \frac{1}{2}\|\left[G_{[s_1+1,s_1+s_2]}, G_{[1,s_1]}\right]\|\delta t^2, \qquad \text{(F.22)}
\end{aligned}$$

where we have used the inductive hypothesis and the Trotter error bound [77, Eq. (143)] in the last inequality. To bound the commutator norm, we use the following lemma.

Lemma 36. *For any* $k_0 \leq k_1, j_0 \leq j_1$, *we have*

$$\|\left[G_{[k_0,k_1]}, G_{[j_0,j_1]}\right]\| \leq 2\left(j_1 - j_0 + k_1 - k_0 + 2\right)\xi\sqrt{m}\|G\|^2. \tag{F.23}$$

Proof. We have

$$\begin{aligned}
\|[G_{k_0\le k\le k_1}, G_{j_0\le j\le j_1}]\| &= \|\left[\sum_{k=k_0}^{k_1} G_k, \sum_{j=j_0}^{j_1} G_j\right]\| \\
&\le \|\left[\sum_{k=k_0}^{k_1} G_k, \sum_{j=j_0}^{j_1} G_j - (j_1-j_0+1)\sum_{\mu=1}^{m} P_\mu G P_\mu\right]\| \\
&\quad + \|\left[\sum_{k=k_0}^{k_1} G_k - (k_1-k_0+1)\sum_{\mu=1}^{m} P_\mu G P_\mu, (j_1-j_0+1)\sum_{\mu=1}^{m} P_\mu G P_\mu\right]\| \\
&\le 2(k_1-k_0+1)\|G\|\left(\xi\sqrt{m}\|G\|\right) + 2\left(\xi\sqrt{m}\|G\|\right)(j_1-j_0+1)\|G\| \\
&= 2(j_1+k_1-j_0-k_0+2)\xi\sqrt{m}\|G\|^2,
\end{aligned}$$

where we have used Corollary 6 to prove the second inequality. Therefore, the lemma follows. □

Applying Lemma 36 to Eq. (F.22), we have

$$\|\prod_{k=1}^{s_1+s_2} e^{-iG_k\delta t} - e^{-iG_{[1,s_1+s_2]}\delta t}\| \le (2s_1\log_2 s_1 + 2s_2\log_2 s_2 + s_1 + s_2)\xi\sqrt{m}\|G\|^2\delta t^2. \tag{F.24}$$

We now apply the above equation repeatedly to prove Lemma 35. Note that Lemma 35 holds trivially for $s = 1$. Suppose that it holds for all $s \le s_0$ for some $s_0 \ge 1$. We shall prove that it holds for $s = s_0 + 1$.

First, we consider the case where s is even, i.e. there exists an integer $l \ge 1$ such

that $s = 2l$. Applying Eq. (F.24) with $s_1 = s_2 = l$, we get

$$\|\prod_{k=1}^{s} e^{-iG_k\delta t} - e^{-iG_{[1,s]}\delta t}\| \leq (2l\log_2 l + 2l\log_2 l + l + l)\xi\sqrt{m}\|G\|^2\delta t^2 \tag{F.25}$$

$$= (2s\log_2(s/2) + s)\xi\sqrt{m}\|G\|^2\delta t^2 \tag{F.26}$$

$$< 2s\log_2 s\ \xi\sqrt{m}\|G\|^2\delta t^2. \tag{F.27}$$

Therefore, Lemma 35 holds if s is even.

When s is odd, there exists an integer $l \geq 1$ such that $s = 2l + 1$. Applying Lemma 35 with $s_1 = l$ and $s_2 = l + 1$, we have

$$\|\prod_{k=1}^{s} e^{-iG_k\delta t} - e^{-iG_{[1,s]}\delta t}\| \leq (2l\log_2 l + 2(l+1)\log_2(l+1) + 2l + 1)\xi\sqrt{m}\|G\|^2\delta t^2. \tag{F.28}$$

Let

$$g(x) = 2x\log_2 x + 2(x+1)\log_2(x+1) + 2x + 1 - 2(2x+1)\log_2(2x+1). \tag{F.29}$$

It is straightforward to verify that $g(1) < 0$ and

$$g'(x) = 2\log_2\frac{2x(1+x)}{(1+2x)^2} < 0 \tag{F.30}$$

for all $x \geq 1$. Therefore, $g(x) \leq 0$ for all $x \geq 1$. Applying this bound to Eq. (F.28), we

get

$$\|\prod_{k=1}^{s} e^{-iG_k\delta t} - e^{-iG_{[1,s]}\delta t}\| \leq 2(2l+1)\log_2(2l+1)\xi\sqrt{m}\|G\|^2\delta t^2 \tag{F.31}$$

$$= 2s\log_2 s\,\xi\sqrt{m}\|G\|^2\delta t^2. \tag{F.32}$$

Thus, Lemma 35 holds for odd s too. By induction, Lemma 35 holds for all $s \geq 1$.

Combining Lemma 35 with

$$\|e^{-i\frac{t}{r}G_{[1,r]}} - e^{-itG_{\text{Zeno}}}\| \leq \frac{t}{r}\|G_{[1,r]} - rG_{\text{Zeno}}\| \leq \frac{\xi\sqrt{m}\|G\|t}{r} \tag{F.33}$$

from Corollary 6, we prove Eq. (F.9).

F.2 Symmetry Protection by the Quantum Zeno Effect

In this section, we make a formal connection between the symmetry protection technique and the quantum Zeno effect. In particular, we show how the quantum Zeno framework provides an alternative explanation for the suppression of simulation error under symmetry protection.

We first note that the symmetry transformations in our scheme are analogous to the kicks in the quantum Zeno framework. Suppose that the symmetry transformations have the form $C_k = C_0^k$, where $C_0 \in \mathcal{S}$ is also a symmetry transformation. Let

$$C_0 = \sum_{\mu} e^{-i\phi_\mu} P_\mu \tag{F.34}$$

be the spectral decomposition of C_0, with $e^{-i\phi_\mu}$ being the distinct eigenvalues and P_μ being the projectors onto the respective eigensubspaces. The condition on $e^{-i\phi_\mu}$ being distinct ensures that C_0 satisfies the definition of U_{kick} in Eq. (F.2).

With $e^{-iH\delta t}$ being approximated by a circuit $S_{\delta t}$ in each time step, our symmetry-protected simulation becomes

$$\prod_{k=1}^{r} C_k^\dagger S_{\delta t} C_k = (C_0^\dagger)^r (e^{-iH_{\text{eff}}\delta t} C_0)^r, \tag{F.35}$$

where H_{eff} is the generator of $S_{\delta t}$ and exists for a small enough δt (see Lemma 10). Comparing Eq. (F.35) with Eq. (F.3), we identify $C_0 = U_{\text{kick}}$. Therefore, by Theorem 13, the symmetry protected simulation is effectively described by

$$\prod_{k=1}^{r} C_k^\dagger S_{\delta t} C_k \to e^{-iH_{\text{eff,Zeno}}t}, \tag{F.36}$$

in the large r limit, where $H_{\text{eff,Zeno}} = \sum_\mu P_\mu H_{\text{eff}} P_\mu$.

Recall that H_{eff} is the effective Hamiltonian corresponding the Trotterized evolution $S_{\delta t}$. For small δt, it is a sum of the true Hamiltonian H that we are simulating and a small error term V (due to the use of Trotterization):

$$H_{\text{eff}} = H + V. \tag{F.37}$$

Therefore, under the symmetry protection, the effective Hamiltonian is replaced by its

projection onto the Zeno subspaces:

$$H_{\text{eff}} \to H_{\text{eff,Zeno}} = H + V_{\text{Zeno}}, \tag{F.38}$$

where $V_{\text{Zeno}} = \sum_\mu P_\mu V P_\mu$ is the corresponding projection of V. In particular, if the error V does not respect the symmetry, the projection V_{Zeno} could be much smaller than the error V in an unprotected simulation. The quantum Zeno framework therefore provides alternative intuition for the error suppression from the symmetry protection. We note, however, that choosing the symmetry transformations C_k independently, instead of $C_k = C_0^k$ considered in this section, could lead to more reduction of the simulation error, and we demonstrate this advantage in Sec. 8.3.

We make these arguments rigorous by proving a bound analogous to that in Theorem 13 for symmetry-protected quantum simulation. Specifically, we consider $G = H_{\text{eff}} = H + V$, where $[H, U_{\text{kick}}] = 0$. Note that under this assumption, the distinctiveness of the eigenvalues of U_{kick} ensures that $[P_\mu, H] = 0$ for all μ in the spectral decomposition of U_{kick}. We will also denote by $V_k = U_{\text{kick}}^{\dagger k} V U_{\text{kick}}^k = G_k - H$.

Theorem 15 (Symmetry protection by quantum Zeno effect)**.** *Let U_{kick} be the unitary defined in Eq.* (F.2) *and suppose that $G = H + V$ such that $[H, U_{kick}] = 0$. Let $G_{Zeno} = \sum_\mu P_\mu G P_\mu = H + \sum_\mu P_\mu V P_\mu$ denote the projection of G onto the subspaces defined by*

a set of orthogonal projectors $\{P_\mu\}$ in the spectral decomposition of U_{kick}. We have

$$\varepsilon_{Zeno} \equiv \|U_{kick}^{\dagger r}\left(U_{kick}e^{-i\frac{t}{r}G}\right)^r - e^{-itG_{Zeno}}\| \leq \frac{2\xi\sqrt{m}\|G\|\|V\|t^2\log r}{r} + \frac{\xi\sqrt{m}\|V\|t}{r} \tag{F.39}$$

$$\leq \frac{3\xi\sqrt{m}\|G\|\|V\|t^2\log r}{r}, \tag{F.40}$$

where ξ is the inverse spectral gap defined in Eq. (F.7).

Note that this bound is stronger than Eq. (F.9) in that the dependence on the norm of the Hamiltonian is improved from $\|G\|^2$ to $\|G\|\|V\|$. To prove Eq. (F.40), we derive a different version of Lemma 35 for the case $G = H + V$, where $[H, U_{\text{kick}}] = 0$.

Lemma 37. *Suppose $G = H + V$, where $[H, U_{kick}] = 0$. For any $s \geq 1$ and δt, we have*

$$\|\prod_{k=1}^{s} e^{-iG_k\delta t} - e^{-iG_{[1,s]}\delta t}\| \leq 2\xi\sqrt{m}\|G\|\|V\|\delta t^2 s\log_2 s. \tag{F.41}$$

Again, we prove Lemma 37 by induction on s. Suppose Lemma 37 holds for $s = s_1$ and $s = s_2$ such that $|s_2 - s_1| \leq 1$, we shall prove that it holds for $s = s_1 + s_2$. Using the triangle inequality

$$\begin{aligned}
&\|\prod_{k=1}^{s_1+s_2} e^{-iG_k\delta t} - e^{-iG_{[1,s_1+s_2]}\delta t}\| \\
&\leq \|\prod_{k=1}^{s_1} e^{-iG_k\delta t} - e^{-iG_{[1,s_1]}\delta t}\| + \|\prod_{k=s_1+1}^{s_1+s_2} e^{-iG_k\delta t} - e^{-iG_{[s_1+1,s_1+s_2]}\delta t}\| \\
&\quad + \|e^{-iG_{[1,s_1+s_2]}\delta t} - e^{-iG_{[s_1+1,s_1+s_2]}\delta t}e^{-iG_{[1,s_1]}\delta t}\| \qquad \text{(F.42)} \\
&\leq 2\xi\sqrt{m}\|G\|\|V\|\delta t^2(s_1\log_2 s_1 + s_2\log_2 s_2) + \frac{1}{2}\|\left[G_{[s_1+1,s_1+s_2]}, G_{[1,s_1]}\right]\|\delta t^2. \qquad \text{(F.43)}
\end{aligned}$$

To bound the commutator norm, we use a modified version of Lemma 36.

Lemma 38. *Given $G = H + V$ and $[H, U_{kick}] = 0$, we have*

$$\|\left[G_{[k_0,k_1]}, G_{[j_0,j_1]}\right]\| \leq 2\left(j_1 - j_0 + k_1 - k_0 + 2\right)\xi\sqrt{m}\|G\|\|V\|. \tag{F.44}$$

Proof. We have

$$\begin{aligned}
&\|[G_{k_0\leq k\leq k_1}, G_{j_0\leq j\leq j_1}]\| = \|\left[\sum_{k=k_0}^{k_1} G_k, \sum_{j=j_0}^{j_1} G_j\right]\| \\
&\leq \|\left[\sum_{k=k_0}^{k_1} G_k, \sum_{j=j_0}^{j_1} G_j - (j_1 - j_0 + 1)\sum_{\mu=1}^{m} P_\mu G P_\mu\right]\| \\
&+ \|\left[\sum_{k=k_0}^{k_1} G_k - (k_1 - k_0 + 1)\sum_{\mu=1}^{m} P_\mu G P_\mu, (j_1 - j_0 + 1)\sum_{\mu=1}^{m} P_\mu G P_\mu\right]\| \\
&= \|\left[\sum_{k=k_0}^{k_1} G_k, \sum_{j=j_0}^{j_1} V_j - (j_1 - j_0 + 1)\sum_{\mu=1}^{m} P_\mu V P_\mu\right]\| \\
&+ \|\left[\sum_{k=k_0}^{k_1} V_k - (k_1 - k_0 + 1)\sum_{\mu=1}^{m} P_\mu V P_\mu, (j_1 - j_0 + 1)\sum_{\mu=1}^{m} P_\mu G P_\mu\right]\| \\
&\leq 2(k_1 - k_0 + 1)\|G\|\left(\xi\sqrt{m}\|V\|\right) + 2\left(\xi\sqrt{m}\|V\|\right)(j_1 - j_0 + 1)\|G\| \\
&= 2(j_1 - j_0 + k_1 - k_0 + 2)\xi\sqrt{m}\|G\|\|V\|,
\end{aligned}$$

where $V_k = U_{\text{kick}}^{\dagger k} V U_{\text{kick}}^{k} = G_k - H$ as mentioned above. Therefore, the lemma follows.

□

Applying Lemma 38 to Eq. (F.43), we have

$$\| \prod_{k=1}^{s_1+s_2} e^{-iG_k\delta t} - e^{-iG_{[1,s_1+s_2]}\delta t} \| \leq (2s_1 \log_2 s_1 + 2s_2 \log_2 s_2 + s_1 + s_2)\xi\sqrt{m}\|G\|\|V\|\delta t^2. \tag{F.45}$$

Using this bound and an inductive argument similar to the proof of Lemma 35, we prove Lemma 37. Finally, combining Lemma 37 at $s = r$ with

$$\|e^{-i\frac{t}{r}G_{[1,r]}} - e^{-itG_{\text{Zeno}}}\| \leq \|G_{[1,r]} - rG_{\text{Zeno}}\|\frac{t}{r} \leq \|\sum_{k=1}^{r} V_k - rV_{\text{Zeno}}\|\frac{t}{r} \leq \frac{\xi\sqrt{m}\|V\|t}{r}, \tag{F.46}$$

we obtain Eq. (F.40).

F.3 A General Bound on the Trotter Error

In Sec. 8.2, we prove a bound on the simulation error under the protection from a special class of symmetry transformations $C_k = C_0^k$. In this section, we prove a similar, but more general, bound without making such an assumption.

Given a fixed total evolution time t, we first estimate the number of Trotter steps r required to simulate $\exp(-iHt)$ so that the total additive error of the simulation meets a threshold ε. Suppose the Hamiltonian $H = \sum_{\mu=1}^{L} H_\mu$ is a sum of L Hamiltonian terms H_μ such that each $e^{-iH_\mu\delta t}$ can be readily simulated on quantum computers. Again, we

define the following quantities

$$\gamma \equiv \sum_{\mu=1}^{L} \sum_{\nu=\mu+1}^{L} \|[H, [H_\nu, H_\mu]]\|, \tag{F.47}$$

$$\beta \equiv \sum_{\mu=1}^{L} \sum_{\nu=\mu+1}^{L} \sum_{\nu'=\nu}^{L} \|[H_{\nu'}, [H_\nu, H_\mu]]\|, \tag{F.48}$$

$$\alpha \equiv \sum_{\mu=1}^{L} \sum_{\nu=\mu+1}^{L} \|[H_\nu, H_\mu]\|, \tag{F.49}$$

that are independent of t, r.

The first-order Trotterization approximates $\exp(-iH\delta t)$ by

$$S_{\delta t} = \prod_{\mu=1}^{L} e^{-iH_\mu \delta t}, \tag{F.50}$$

where $\prod_{\mu=1}^{L} U_\mu \equiv U_L \ldots U_2 . U_1$ is an ordered product.

To get an accurate scaling of the gate count with the error tolerance, time, and the system size, we extend the approach in Ref. [77] to estimate the higher-order contributions to the total error. First, we estimate the higher-order contributions to the additive error in one Trotter step.

Lemma 39. *Assuming $\beta\delta t \leq 2\alpha$ and $\alpha^2\delta t \leq \gamma + \beta$, the Trotter error in approximating $U_{\delta t} = \exp(-iH\delta t)$ by $S_{\delta t}$ in Eq. (F.50) is given by*

$$\mathcal{E}_{\delta t} \equiv U_{\delta t} - S_{\delta t} = U_{\delta t} v_0 \frac{\delta t^2}{2} + \widetilde{\mathcal{V}}(\delta t), \tag{F.51}$$

where v_0 is defined in Eq. (8.15) and $\widetilde{\mathcal{V}}(\delta t)$ is an operator bounded by

$$\|\widetilde{\mathcal{V}}(\delta t)\| \leq \Lambda \delta t^3, \tag{F.52}$$

with $\Lambda = \frac{5}{6}(\gamma + \beta)$.

Proof. From [77, Theorem 8], we have

$$S_{\delta t} = e^{-iH\delta t}\mathcal{T}\exp\left\{-i\int_0^{\delta t} d\tau_1 \widetilde{F}(\tau_1)\right\}, \tag{F.53}$$

where $\mathcal{T}\exp\{\}$ is the time-ordered exponential,

$$\widetilde{F}(\tau_1) = e^{i\tau_1 \mathrm{ad}_H} \sum_{\mu=1}^{L}\left(\prod_{\nu=\mu+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_\nu}} H_\mu - H_\mu\right), \tag{F.54}$$

$\mathrm{ad}_A B \equiv [A, B]$, and $e^{-it\mathrm{ad}_A} B = e^{-itA} B e^{itA}$. Note that the summand in the definition of $\widetilde{F}(\tau_1)$ is of order $\mathcal{O}(\tau_1)$. Therefore, we can rewrite it as (See [77, Theorem 10] or use a direct differentiation):

$$\prod_{\nu=\mu+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_\nu}} H_\mu - H_\mu = -i\sum_{\nu=\mu+1}^{L}\int_0^{\tau_1} d\tau_2 \prod_{\nu'=\nu+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_{\nu'}}} e^{-i\tau_2 \mathrm{ad}_{H_\nu}} [H_\nu, H_\mu] \tag{F.55}$$

$$= -i\sum_{\nu=\mu+1}^{L}[H_\nu, H_\mu]\tau_1 - i\sum_{\nu=\mu+1}^{L}\int_0^{\tau_1} d\tau_2 \underbrace{\left(\prod_{\nu'=\nu+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_{\nu'}}} e^{-i\tau_2 \mathrm{ad}_{H_\nu}} [H_\nu, H_\mu] - [H_\nu, H_\mu]\right)}_{\equiv G_{\mu,\nu}(\tau_1,\tau_2)}. \tag{F.56}$$

Again, we note that $G(\tau_1) = \mathcal{O}(\tau_1 + \tau_2)$. Therefore, we can rewrite it (using either [77]

or a direct differentiation) as

$$G_{\mu,\nu}(\tau_1,\tau_2) = -i\sum_{\nu'=\nu+1}^{L}\prod_{s=\nu'+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_s}}\int_0^{\tau_1} d\tau_3 e^{-i\tau_3 \mathrm{ad}_{H_{\nu'}}}\left[H_{\nu'},[H_\nu,H_\mu]\right] \\ -i\frac{\tau_2}{\tau_1}\prod_{\nu'=\nu+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_{\nu'}}}\int_0^{\tau_1} d\tau_3 e^{-i\tau_3\frac{\tau_2}{\tau_1}\mathrm{ad}_{H_\nu}}\left[H_\nu,[H_\nu,H_\mu]\right]. \tag{F.57}$$

Using the triangle inequality, we have

$$\|G_{\mu,\nu}(\tau_1,\tau_2)\| \leq \tau_1\sum_{\nu'=\nu}^{L}\|[H_{\nu'},[H_\nu,H_\mu]]\|. \tag{F.58}$$

Therefore,

$$\|\widetilde{F}(\tau_1)\| \leq \tau_1\sum_{\mu=1}^{L}\sum_{\nu=\mu+1}^{L}\|[H_\nu,H_\mu]\| + \frac{\tau_1^2}{2}\sum_{\mu=1}^{L}\sum_{\nu=\mu+1}^{L}\sum_{\nu'=\nu}^{L}\|[H_{\nu'},[H_\nu,H_\mu]]\| \tag{F.59}$$

In addition, we have

$$e^{i\tau_1 \mathrm{ad}_H}\left[H_\nu,H_\mu\right] - \left[H_\nu,H_\mu\right] = i\int_0^{\tau_1} d\tau_2 e^{i\tau_2 \mathrm{ad}_H}\left[H,[H_\nu,H_\mu]\right]. \tag{F.60}$$

Therefore, with $v_0 = \sum_{\mu=1}^{L}\sum_{\nu=\mu+1}^{L}\left[H_\nu,H_\mu\right]$, we have

$$\begin{aligned}\widetilde{F}(\tau_1) + iv_0\tau_1 &= -i\sum_{\mu=1}^{L}\sum_{\nu=\mu+1}^{L}\left(e^{i\tau_1 \mathrm{ad}_H}\left[H_\nu,H_\mu\right]\tau_1 - \left[H_\nu,H_\mu\right]\tau_1 + e^{i\tau_1 \mathrm{ad}_H}\int_0^{\tau_1} d\tau_2 G_{\mu,\nu}(\tau_1,\tau_2)\right) \\ &= -i\sum_{\mu=1}^{L}\sum_{\nu=\mu+1}^{L}\int_0^{\tau_1} d\tau_2\left(i\tau_1 e^{i\tau_2 \mathrm{ad}_H}\left[H,[H_\nu,H_\mu]\right] + e^{i\tau_1 H}G_{\mu,\nu}(\tau_1,\tau_2)\right)\end{aligned} \tag{F.61}$$

Expanding the time-ordered exponential, we have

$$\begin{aligned}
&\mathcal{T}\exp\left\{-i\int_0^{\delta t} d\tau_1 \widetilde{F}(\tau_1)\right\} \\
&= \mathbb{I} - i\int_0^{\delta t} d\tau_1 \widetilde{F}(\tau_1) - \int_0^{\delta t} d\tau_1 \int_0^{\tau_1} d\tau_2 \widetilde{F}(\tau_1)\widetilde{F}(\tau_2)\mathcal{T}\exp\left\{-i\int_0^{\tau_2} d\tau_3 \widetilde{F}(\tau_3)\right\}
\end{aligned} \tag{F.62}$$

$$\begin{aligned}
&= \mathbb{I} - \frac{\delta t^2}{2} v_0 - i\int_0^{\delta t} d\tau_1 [\widetilde{F}(\tau_1) + i v_0 \tau_1] \\
&\quad - \int_0^{\delta t} d\tau_1 \int_0^{\tau_1} d\tau_2 \widetilde{F}(\tau_1)\widetilde{F}(\tau_2)\mathcal{T}\exp\left\{-i\int_0^{\tau_2} d\tau_3 \widetilde{F}(\tau_3)\right\}.
\end{aligned} \tag{F.63}$$

Therefore, we have

$$\begin{aligned}
&\|S_{\delta t} - e^{-iH\delta t} + e^{-iH\delta t} v_0 \frac{\delta t^2}{2}\| \\
&\leq \int_0^{\delta t} d\tau_1 \|\widetilde{F}(\tau_1) + i v_0 \tau_1\| + \int_0^{\delta t} d\tau_1 \int_0^{\tau_1} d\tau_2 \|\widetilde{F}(\tau_1)\| \|\widetilde{F}(\tau_2)\| \qquad \text{(F.64)} \\
&\leq \frac{\gamma+\beta}{3}\delta t^3 + \frac{\delta t^4}{8}\left(\alpha + \frac{\delta t}{2}\beta\right)^2. \qquad \text{(F.65)}
\end{aligned}$$

In particular, assuming $\beta\delta t \leq 2\alpha$ and $\alpha^2\delta t \leq \gamma + \beta$, we have

$$\|\mathcal{E}_{\delta t} - U_{\delta t} v_0 \frac{\delta t^2}{2}\| \leq \Lambda \delta t^3, \tag{F.66}$$

with $\Lambda = \frac{5}{6}(\gamma + \beta)$. Therefore, Lemma 39 follows. □

As a result of Lemma 39, we can bound the additive error in one Trotter step:

$$\|\mathcal{E}_{\delta t}\| \leq \frac{\|v_0\|}{2}\delta t^2 + \Lambda \delta t^3. \tag{F.67}$$

Therefore, we arrive at a bound for the total error for the simulation

$$\varepsilon = \|U_t - \prod_{k=1}^{r} C_k^\dagger S_{\delta t} C_k\| \tag{F.68}$$

$$\leq \|\sum_{k=1}^{r} C_k^\dagger U_{k\delta t}^\dagger \mathcal{E}_{\delta t} U_{k\delta t} C_k\| + \sum_{j=2}^{r} \binom{r}{j} \|\mathcal{E}_{\delta t}\|^j \tag{F.69}$$

$$\leq \|\sum_{k=1}^{r} C_k^\dagger U_{k\delta t}^\dagger v_0 U_{k\delta t} C_k\| \frac{\delta t^2}{2} + r\Lambda\delta t^3 + \sum_{j=2}^{r} (r\|\mathcal{E}_{\delta t}\|)^j \tag{F.70}$$

$$\leq \|\sum_{k=1}^{r} C_k^\dagger U_{k\delta t}^\dagger v_0 U_{k\delta t} C_k\| \frac{\delta t^2}{2} + r\Lambda\delta t^3 + 2r^2\|\mathcal{E}_{\delta t}\|^2 \tag{F.71}$$

$$\leq \Big\| \underbrace{\frac{1}{r}\sum_{k=1}^{r} C_k^\dagger U_{k\delta t}^\dagger v_0 U_{k\delta t} C_k}_{\equiv \overline{v_0}} \Big\| \frac{t^2}{2r} + \Lambda\frac{t^3}{r^2} + 2\left(\frac{1}{2}\|v_0\| + \Lambda\frac{t}{r}\right)^2 \frac{t^4}{r^2}, \tag{F.72}$$

where $U_{k\delta t} = \exp(-iHk\delta t)$ and we have assume $r\|\mathcal{E}_{\delta t}\| \leq 1/2$ to bound the sum over j. This bound again has the same feature as the bound in Theorem 10: the total error, to the lowest-order, scales with $\|\overline{v_0}\|$—an averaged version of v_0 under the symmetry transformations—instead of scaling with $\|v_0\|$. Note, however, that the definition of $\overline{v_0}$ here, with the addition of the transformations under $U_{k\delta t}$, is slightly different from Theorem 10.

F.4 Proof of Lemma 10

In this section, we prove Lemma 10, which provides a bound on the error in one Trotter step.

Proof. From [77, Theorem 8], we have

$$S_{\delta t} = \mathcal{T} \exp\left\{-i \int_0^{\delta t} d\tau_1 \left(H + F(\tau_1)\right)\right\}, \tag{F.73}$$

where $\mathcal{T} \exp\{\}$ is the time-ordered exponential,

$$F(\tau_1) = \sum_{\mu=1}^{L} \left(\prod_{\nu=\mu+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_\nu}} H_\mu - H_\mu \right), \tag{F.74}$$

$\mathrm{ad}_A B \equiv [A, B]$, and $e^{-it\mathrm{ad}_A} B = e^{-itA} B e^{itA}$. Note that the summand in the definition of $F(\tau_1)$ is of order $\mathcal{O}(\tau_1)$. Therefore, we can rewrite it as (See [77, Theorem 10] or use a direct differentiation):

$$\prod_{\nu=\mu+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_\nu}} H_\mu - H_\mu = -i \sum_{\nu=\mu+1}^{L} \int_0^{\tau_1} d\tau_2 \prod_{\nu'=\nu+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_{\nu'}}} e^{-i\tau_2 \mathrm{ad}_{H_\nu}} [H_\nu, H_\mu] \tag{F.75}$$

$$= -i \sum_{\nu=\mu+1}^{L} [H_\nu, H_\mu] \tau_1 - i \sum_{\nu=\mu+1}^{L} \int_0^{\tau_1} d\tau_2 \underbrace{\left(\prod_{\nu'=\nu+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_{\nu'}}} e^{-i\tau_2 \mathrm{ad}_{H_\nu}} [H_\nu, H_\mu] - [H_\nu, H_\mu] \right)}_{\equiv G_{\mu,\nu}(\tau_1, \tau_2)}. \tag{F.76}$$

We note that $G(\tau_1) = \mathcal{O}(\tau_1 + \tau_2)$ [Recall that $\mathcal{O}()$ is the standard Bachmann-Landau big-O notation.] Therefore, we can rewrite it (using either [77] or a direct differentiation) as

$$G_{\mu,\nu}(\tau_1, \tau_2) = -i \sum_{\nu'=\nu+1}^{L} \prod_{s=\nu'+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_s}} \int_0^{\tau_1} d\tau_3 e^{-i\tau_3 \mathrm{ad}_{H_{\nu'}}} [H_{\nu'}, [H_\nu, H_\mu]]$$
$$-i \frac{\tau_2}{\tau_1} \prod_{\nu'=\nu+1}^{L} e^{-i\tau_1 \mathrm{ad}_{H_{\nu'}}} \int_0^{\tau_1} d\tau_3 e^{-i\tau_3 \frac{\tau_2}{\tau_1} \mathrm{ad}_{H_\nu}} [H_\nu, [H_\nu, H_\mu]]. \tag{F.77}$$

Using the triangle inequality, we have

$$\|G_{\mu,\nu}(\tau_1,\tau_2)\| \leq \tau_1 \sum_{\nu'=\nu}^{L} \|[H_{\nu'},[H_\nu,H_\mu]]\|. \tag{F.78}$$

Therefore, with $v_0 = \sum_{\mu=1}^{L}\sum_{\nu=\mu+1}^{L}[H_\nu,H_\mu]$, we have

$$F(\tau_1) + iv_0\tau_1 = \underbrace{-i\sum_{\mu=1}^{L}\sum_{\nu=\mu+1}^{L}\int_0^{\tau_1} d\tau_2 G_{\mu,\nu}(\tau_1,\tau_2)}_{\equiv\mathcal{F}(\tau_1)}. \tag{F.79}$$

Using the bound on $\|G_{\mu,\nu}\|$, we have

$$\|\mathcal{F}(\tau_1)\| \leq \tau_1^2 \sum_{\mu=1}^{L}\sum_{\nu=\mu+1}^{L}\sum_{\nu'=\nu}^{L}\|[H_{\nu'},[H_\nu,H_\mu]]\|, \tag{F.80}$$

which implies

$$\|F(\tau_1)\| \leq \tau_1 \underbrace{\sum_{\mu=1}^{L}\sum_{\nu=\mu+1}^{L}\|[H_\nu,H_\mu]\|}_{\equiv\alpha} + \tau_1^2 \underbrace{\sum_{\mu=1}^{L}\sum_{\nu=\mu+1}^{L}\sum_{\nu'=\nu}^{L}\|[H_{\nu'},[H_\nu,H_\mu]]\|}_{\equiv\beta}. \tag{F.81}$$

In addition, combining Eq. (F.79) with Eq. (F.73), we have

$$S_{\delta t} = \mathcal{T}\exp\left\{-i\int_0^{\delta t} d\tau_1\left[H - iv_0\tau_1 + \mathcal{F}(\tau_1)\right]\right\}, \tag{F.82}$$

with v_0 and $\mathcal{F}(\tau_1)$ given above.

Next, we rewrite the time-ordered exponential into a regular exponential using the Magnus expansion.

Lemma 40 (Magnus expansion [152, 205, 206]). *Let $\mathcal{A}(\tau)$ be a continuous operator-valued function defined for $0 \leq \tau \leq t$ such that $\int_0^t \mathrm{d}\tau \|\mathcal{A}(\tau)\| \leq 1$. Then, the equality*

$$\mathcal{T} \exp\left\{\int_0^t \mathrm{d}\tau\, \mathcal{A}(\tau)\right\} = \exp\left(\sum_{j=1}^{\infty} \Omega_j(t)\right) \tag{F.83}$$

holds with a convergent operator series $\sum_{j=1}^{\infty} \Omega_j(t)$, where

$$\Omega_j(t) = \frac{1}{j^2} \sum_{\sigma} (-1)^{d_b} \frac{1}{\binom{j-1}{d_b}} \int_0^t d\tau_1 \cdots \int_0^{d\tau_{j-1}} d\tau_j \left[\mathcal{A}(\tau_1), \ldots [\mathcal{A}(\tau_{j-1}), \mathcal{A}(\tau_j)] \ldots\right], \tag{F.84}$$

with the sum being taken over all permutations σ of $\{1, \ldots, j\}$ and d_b is the number of descents, i.e. pairs of consecutive numbers σ_k, σ_{k+1} for $k = 1, \ldots, j-1$ such that $\sigma_k > \sigma_{k+1}$, in the permutation σ. Furthermore, $\Omega_j(t)$ are all anti-Hermitian if $\mathcal{A}(\tau)$ is anti-Hermitian. It is worth noting that the first two $\Omega_j(t)$ are

$$\Omega_1(t) = \int_0^t \mathrm{d}\tau\, \mathcal{A}(\tau), \tag{F.85}$$

$$\Omega_2(t) = \frac{1}{2} \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 \left[\mathcal{A}(\tau_1), \mathcal{A}(\tau_2)\right]. \tag{F.86}$$

We now use Lemma 40 to rewrite Eq. (F.82) with $\mathcal{A}(\tau) = -i[H + F(\tau)] = -i[H - iv_0\tau_1 + \mathcal{F}(\tau_1)]$:

$$S_{\delta t} = \exp\left(\sum_{j=1}^{\infty} \Omega_j(\delta t)\right), \tag{F.87}$$

where the first-order Magnus term is

$$\Omega_1(\delta t) = -i \int_0^{\delta t} d\tau_1 \left[H - i v_0 \tau_1 + \mathcal{F}(\tau_1) \right] = -i\delta t \left[H - \frac{i}{2} v_0 \delta t + \frac{1}{\delta t} \int_0^{\delta t} d\tau_1 \mathcal{F}(\tau_1) \right]. \tag{F.88}$$

To bound the higher-order terms in the Magnus expansion, we first note that

$$\begin{aligned}
\| [\mathcal{A}(\tau_1), \mathcal{A}(\tau_2)] \| &= \| [H + F(\tau_1), H + F(\tau_2)] \| && \text{(F.89)} \\
&\leq 2 \Big(\|H\| \|F(\tau_2)\| + \|F(\tau_1)\| \|H\| + \|F(\tau_1)\| \|F(\tau_2)\| \Big) && \text{(F.90)} \\
&\leq 2 \Big(\|H\| \left(\alpha\tau_1 + \beta\tau_1^2 \right) + \|H\| \left(\alpha\tau_2 + \beta\tau_2^2 \right) + \left(\alpha\tau_1 + \beta\tau_1^2 \right) \left(\alpha\tau_2 + \beta\tau_2^2 \right) \Big) && \text{(F.91)} \\
&\leq 2 \left(2\|H\| + \alpha\delta t + \beta\delta t^2 \right) \left(\alpha\delta t + \beta\delta t^2 \right) && \text{(F.92)} \\
&\leq 4 \left(\|H\| + \alpha\delta t + \beta\delta t^2 \right) \left(\alpha\delta t + \beta\delta t^2 \right) && \text{(F.93)}
\end{aligned}$$

for all $\tau_1, \tau_2 \leq \delta t$. Similarly, for higher-order nested commutators:

$$\begin{aligned}
\| [\mathcal{A}(\tau_1), \ldots [\mathcal{A}(\tau_{j-1}), \mathcal{A}(\tau_j)] \ldots] \| &\leq 2^{j-2} \|\mathcal{A}(\tau_1)\| \ldots \|\mathcal{A}(\tau_{j-2})\| \| [\mathcal{A}(\tau_{j-1}), \mathcal{A}(\tau_j)] \| && \text{(F.94)} \\
&\leq 2^j \left(\|H\| + \alpha\delta t + \beta\delta t^2 \right)^{j-1} \left(\alpha\delta t + \beta\delta t^2 \right). && \text{(F.95)}
\end{aligned}$$

Using Lemma 40 and noting that there are $j!$ permutations for each j, we can crudely

bound

$$\|\Omega_j(\delta t)\| \leq \sum_\sigma \int_0^{\delta t} d\tau_1 \int_0^{d\tau_{j-1}} d\tau_j \|[\mathcal{A}(\tau_1), \dots [\mathcal{A}(\tau_{j-1}), \mathcal{A}(\tau_j)] \dots]\| \tag{F.96}$$

$$\leq j! \frac{\delta t^j}{j!} 2^j \left(\|H\| + \alpha\delta t + \beta\delta t^2\right)^{j-1} \left(\alpha\delta t + \beta\delta t^2\right) \tag{F.97}$$

$$\leq (2\delta t)^j \left(\|H\| + \alpha\delta t + \beta\delta t^2\right)^{j-1} \left(\alpha\delta t + \beta\delta t^2\right) \tag{F.98}$$

for all $j \geq 2$. Define

$$\mathcal{V}(\delta t) \equiv \frac{1}{\delta t} \int_0^{\delta t} d\tau_1 \mathcal{F}(\tau_1) + \frac{1}{\delta t} \sum_{j=2}^{\infty} \Omega_j(\delta t), \tag{F.99}$$

we could write

$$S_{\delta t} = \exp\left[-i\delta t \big(H \underbrace{-\frac{i}{2} v_0 \delta t + \mathcal{V}(\delta t)}_{\equiv V} \big) \right]. \tag{F.100}$$

It follows from the bounds on Ω_j above that

$$\|\mathcal{V}(\delta t)\| \leq \delta t^2 \beta + (\alpha + \beta\delta t) \sum_{j=2}^{\infty} (2\delta t)^j \left(\|H\| + \alpha\delta t + \beta\delta t^2\right)^{j-1} \tag{F.101}$$

$$\leq \delta t^2 \beta + 4\delta t^2 (\alpha + \beta\delta t) \left(\|H\| + \alpha\delta t + \beta\delta t^2\right) \sum_{j=0}^{\infty} (2\delta t)^j \left(\|H\| + \alpha\delta t + \beta\delta t^2\right)^j$$

(F.102)

$$\leq \delta t^2 \beta + 8\delta t^2 (\alpha + \beta\delta t) \left(\|H\| + \alpha\delta t + \beta\delta t^2\right), \tag{F.103}$$

where we have assumed $\delta t\big(\|H\| + \alpha\delta t + \beta\delta t^2\big) \leq 1/4$ so that the sum over j in the second

line converges. We note that this assumption also ensures that the Magnus expansion converges. The bound states that $\|\mathcal{V}(\delta t)\|$ scales with δt as $\mathcal{O}\left(\delta t^2\right)$.

Assuming $\beta\delta t \leq \alpha$, $2\alpha\delta t \leq \|H\|$, and $8\delta t\|H\| \leq 1$, we get

$$\|\mathcal{V}(\delta t)\| \leq \delta t^2 \left(\beta + 32\alpha\|H\|\right). \tag{F.104}$$

This bound completes the proof of Lemma 10. Note that the constant prefactor of our bound may be further tightened by using a stronger version of Lemma 40. Such an improvement may be especially useful for near-term implementations of quantum simulation, but a detailed discussion falls out of the scope of the current dissertation and will be left as a subject for future investigation. □

F.5 Proof of Theorem 10

In this section, we provide more details on the proof of Theorem 10 for completeness. Using the triangle inequality

$$\begin{aligned}
\varepsilon &= \|\prod_{k=1}^{r} e^{-i(H+C_k V C_k)\delta t} - e^{-iHt}\| \\
&\leq \|e^{-i\overline{H}_{\text{eff}}t} - e^{-iHt}\| + \|\prod_{k=1}^{r} e^{-i(H+C_k V C_k)\delta t} - e^{-i\overline{H}_{\text{eff}}t}\| && \text{(F.105)} \\
&\leq \|\overline{V}\|t + \frac{2\xi\sqrt{m}(\|H\| + \|V\|)\|V\|t^2 \log r}{r} && \text{(F.106)} \\
&\leq \frac{t^2}{2r}\|\overline{v_0}\| + \chi\frac{t^3}{r^2} + \frac{2\xi\sqrt{m}(\|H\| + \frac{1}{2}\alpha\delta t + \chi\delta t^2)(\frac{1}{2}\alpha\delta t + \chi\delta t^2)t^2 \log r}{r}. && \text{(F.107)}
\end{aligned}$$

Since $\chi = \beta + 32\alpha\|H\|$, $\chi\delta t = \beta\delta t + 32\alpha\|H\|\delta t \leq 5\alpha$ (assuming $\beta\delta t \leq \alpha$ and $8\|H\|\delta t \leq 1$). Therefore, we could upper bound

$$(\|H\| + \frac{1}{2}\alpha\delta t + \chi\delta t^2)(\frac{1}{2}\alpha\delta t + \chi\delta t^2) < 6(\|H\| + 6\alpha\delta t)\alpha\delta t \leq 24\|H\|\alpha\frac{t}{r}, \tag{F.108}$$

where we have also used the assumption that $2\alpha\delta t \leq \|H\|$. Therefore, we have

$$\varepsilon \leq \frac{t^2}{2r}\|\overline{v_0}\| + \chi\frac{t^3}{r^2} + \underbrace{48\xi\sqrt{m}\alpha\|H\|}_{\equiv\kappa}\frac{t^3\log r}{r^2}. \tag{F.109}$$

This completes the proof of Theorem 10.

Bibliography

[1] Seth Lloyd. Universal Quantum Simulators. *Science*, 273(5278):1073–1078, 1996.

[2] Masuo Suzuki. Decomposition formulas of exponential operators and lie exponentials with some applications to quantum mechanics and statistical physics. *J. Math. Phys.*, 26(4):601–612, 1985.

[3] J Huyghebaert and H De Raedt. Product formula methods for time-dependent schrodinger problems. *J. Phys. A: Math. Gen.*, 23(24):5777–5793, dec 1990.

[4] Andrew M. Childs and Yuan Su. Nearly optimal lattice simulation by product formulas. *Phys. Rev. Lett.*, 123:050503, Aug 2019.

[5] Andrew M. Childs. *Quantum information processing in continuous time*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, June 2004.

[6] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders. Efficient Quantum Algorithms for Simulating Sparse Hamiltonians. *Communications in Mathematical Physics*, 270:359–371, March 2007.

[7] Andrew M. Childs, Dmitri Maslov, Yunseong Nam, Neil J. Ross, and Yuan Su. Toward the first quantum simulation with quantum speedup. *Proc. Natl. Acad. Sci.*, 115(38):9456–9461, 2018.

[8] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Simulating Hamiltonian dynamics with a truncated Taylor series. *Phys. Rev. Lett.*, 114(9):090502, March 2015.

[9] Guang Hao Low, Vadym Kliuchnikov, and Nathan Wiebe. Well-conditioned multiproduct Hamiltonian simulation. *arXiv:1907.11679 [physics, physics:quant-ph]*, September 2019.

[10] Guang Hao Low and Isaac L. Chuang. Optimal Hamiltonian Simulation by Quantum Signal Processing. *Phys. Rev. Lett.*, 118:010501, Jan 2017.

[11] Guang Hao Low and Isaac L. Chuang. Hamiltonian Simulation by Qubitization. *Quantum*, 3:163, July 2019.

[12] Elliott H. Lieb and Derek W. Robinson. The finite group velocity of quantum spin systems. *Commun. Math. Phys.*, 28(3):251–257, 1972.

[13] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen. Spin squeezing and reduced quantum noise in spectroscopy. *Phys. Rev. A*, 46(11):R6797–R6800, December 1992.

[14] Michael Foss-Feig, Zhe-Xuan Gong, Alexey V. Gorshkov, and Charles W. Clark. Entanglement and spin-squeezing without infinite-range interactions, 2016.

[15] Norbert M. Linke, Dmitri Maslov, Martin Roetteler, Shantanu Debnath, Caroline Figgatt, Kevin A. Landsman, Kenneth Wright, and Christopher Monroe. Experimental comparison of two quantum computing architectures. *Proceedings of the National Academy of Sciences*, 114(13):3305–3310, March 2017.

[16] Abhinav Deshpande, Bill Fefferman, Minh C. Tran, Michael Foss-Feig, and Alexey V. Gorshkov. Dynamical Phase Transitions in Sampling Complexity. *Phys. Rev. Lett.*, 121(3):030501, 2018.

[17] K. A. Landsman, C. Figgatt, T. Schuster, N. M. Linke, B. Yoshida, N. Y. Yao, and C. Monroe. Verified quantum information scrambling. *Nature*, 567(7746):61–65, March 2019.

[18] Matthew B. Hastings and Tohru Koma. Spectral Gap and Exponential Decay of Correlations. *Commun. Math. Phys.*, 265(3):781–804, August 2006.

[19] Michael Foss-Feig, Zhe-Xuan Gong, Charles W. Clark, and Alexey V. Gorshkov. Nearly Linear Light Cones in Long-Range Interacting Quantum Systems. *PRL*, 114:157201, Apr 2015.

[20] Dominic V. Else, Francisco Machado, Chetan Nayak, and Norman Y. Yao. Improved lieb-robinson bound for many-body hamiltonians with power-law interactions. *Phys. Rev. A*, 101:022333, Feb 2020.

[21] Minh C. Tran, Andrew Y. Guo, Yuan Su, James R. Garrison, Zachary Eldredge, Michael Foss-Feig, Andrew M. Childs, and Alexey V. Gorshkov. Locality and digital quantum simulation of power-law interactions. *Phys. Rev. X*, 9(3):031006, July 2019.

[22] Chi-Fang Chen and Andrew Lucas. Finite speed of quantum scrambling with long range interactions. *Phys. Rev. Lett.*, 123(25):250605, December 2019.

[23] Tomotaka Kuwahara and Keiji Saito. Strictly linear light cones in long-range interacting systems of arbitrary dimensions. *Phys. Rev. X*, 10:031010, Jul 2020.

[24] Zachary Eldredge, Zhe-Xuan Gong, Jeremy T. Young, Ali Hamed Moosavian, Michael Foss-Feig, and Alexey V. Gorshkov. Fast Quantum State Transfer and Entanglement Renormalization Using Long-Range Interactions. *Phys. Rev. Lett.*, 119:170503, Oct 2017.

[25] Minh C. Tran, Chi-Fang Chen, Adam Ehrenberg, Andrew Y. Guo, Abhinav Deshpande, Yifan Hong, Zhe-Xuan Gong, Alexey V. Gorshkov, and Andrew Lucas. Hierarchy of Linear Light Cones with Long-Range Interactions. *Phys. Rev. X*, 10(3):031009, July 2020.

[26] Minh C. Tran, Abhinav Deshpande, Andrew Y. Guo, Andrew Lucas, and Alexey V. Gorshkov. Optimal State Transfer and Entanglement Generation in Power-law Interacting Systems, 2020.

[27] Bruno Nachtergaele, Yoshiko Ogata, and Robert Sims. Propagation of Correlations in Quantum Lattice Systems. *Journal of Statistical Physics*, 124(1):1–13, Jul 2006.

[28] Bruno Nachtergaele and Robert Sims. Lieb-Robinson Bounds and the Exponential Clustering Theorem. *Communications in Mathematical Physics*, 265(1):119–130, Jul 2006.

[29] Zhe-Xuan Gong, Michael Foss-Feig, Spyridon Michalakis, and Alexey V. Gorshkov. Persistence of Locality in Systems with Power-Law Interactions. *PRL*, 113:030602, Jul 2014.

[30] David-Maximilian Storch, Mauritz van den Worm, and Michael Kastner. Interplay of soundcone and supersonic propagation in lattice models with power law interactions. *New J. Phys.*, 17(6):063021, June 2015.

[31] Bruno Nachtergaele, Hillel Raz, Benjamin Schlein, and Robert Sims. Lieb-robinson bounds for harmonic and anharmonic lattice systems. *Communications in Mathematical Physics*, 286(3):1073–1098, Mar 2009.

[32] Isabeau Prémont-Schwarz, Alioscia Hamma, Israel Klich, and Fotini Markopoulou-Kalamara. Lieb-robinson bounds for commutator-bounded operators. *Phys. Rev. A*, 81:040102, Apr 2010.

[33] Isabeau Prémont-Schwarz and Jeff Hnybida. Lieb-robinson bounds on the speed of information propagation. *Phys. Rev. A*, 81:062107, Jun 2010.

[34] Chi-Fang Chen and Andrew Lucas. Operator growth bounds from graph theory. *arXiv:1905.03682 [hep-th, physics:math-ph, physics:quant-ph]*, May 2019.

[35] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu. Theory of trotter error with commutator scaling. *Phys. Rev. X*, 11:011020, Feb 2021.

[36] S. Bravyi, M. B. Hastings, and F. Verstraete. Lieb-Robinson bounds and the generation of correlations and topological quantum order. *arXiv:quant-ph/0603121*, March 2006.

[37] Stefano Chessa, Marco Fanizza, and Vittorio Giovannetti. Quantum-capacity bounds in spin-network communication channels. *Phys. Rev. A*, 100:032311, Sep 2019.

[38] Tomotaka Kuwahara and Keiji Saito. Polynomial growth of out-of-time-order correlator in arbitrary realistic long-range interacting systems. *arXiv:2009.10124*, September 2020.

[39] Chi-Fang Chen. Concentration of otoc and lieb-robinson velocity in random hamiltonians, 2021.

[40] J. J . Bollinger, Wayne M. Itano, D. J. Wineland, and D. J. Heinzen. Optimal frequency measurements with maximally correlated states. *Phys. Rev. A*, 54(6):R4649–R4652, December 1996.

[41] S. Wehner, D. Elkouss, and R. Hanson. Quantum internet: A vision for the road ahead. *Science*, 362(6412), October 2018.

[42] R. P. Feynman. Simulating physics with computers. *Int. J. Theor. Phys.*, 21(6):467–488, June 1982.

[43] P. W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, November 1994.

[44] P. Pham and K. M. Svore. A 2d nearest-neighbor quantum architecture for factoring in polylogarithmic depth. *Quantum Info. Comput.*, 13:937, 2013.

[45] A. Y. Guo, A. Deshpande, S.-K. Chu, Z. Eldredge, P. Bienias, D. Devulapalli, Y. Su, A. M. Childs, and A. V. Gorshkov. Implementing a Fast Unbounded Quantum Fanout Gate Using Power-Law Interactions. *arXiv:2007.00662 [quant-ph]*, July 2020.

[46] M. Christandl and S. Wehner. Quantum anonymous transmissions. In Bimal Roy, editor, *Advances in Cryptology - ASIACRYPT 2005*, pages 217–235, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[47] G. Brassard, A. Broadbent, J. Fitzsimons, S. Gambs, and A. Tapp. Anonymous quantum communication. In Yvo Desmedt, editor, *Information Theoretic Security*, pages 181–182, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[48] M. Hillery, V. Bužek, and A. Berthiaume. Quantum secret sharing. *Phys. Rev. A*, 59(3):1829–1834, March 1999.

[49] K. Kim, S. Korenblit, R. Islam, E. E. Edwards, M.-S. Chang, C. Noh, H. Carmichael, G.-D. Lin, L.-M. Duan, C. C. Joseph Wang, J. K. Freericks, and C. Monroe. Quantum simulation of the transverse Ising model with trapped ions. *New J. Phys.*, 13(10):105003, October 2011.

[50] Joseph W. Britton, Brian C. Sawyer, Adam C. Keith, C.-C. Joseph Wang, James K. Freericks, Hermann Uys, Michael J. Biercuk, and John J. Bollinger. Engineered two-dimensional Ising interactions in a trapped-ion quantum simulator with hundreds of spins. *Nature*, 484(7395):489–492, April 2012.

[51] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe. Demonstration of a small programmable quantum computer with atomic qubits. *Nature*, 536(7614):63–66, August 2016.

[52] Guido Pagano, Aniruddha Bapat, Patrick Becker, Katherine S. Collins, Arinjoy De, Paul W. Hess, Harvey B. Kaplan, Antonis Kyprianidis, Wen Lin Tan, Christopher Baldwin, Lucas T. Brady, Abhinav Deshpande, Fangli Liu, Stephen Jordan, Alexey V. Gorshkov, and Christopher Monroe. Quantum approximate optimization of the long-range ising model with a trapped-ion quantum simulator. *Proc. Natl. Acad. Sci.*, 117(41):25396–25401, 2020.

[53] J. S. Douglas, H. Habibian, C.-L. Hung, A. v. Gorshkov, H. J. Kimble, and D. E. Chang. Quantum Many-Body Models With Cold Atoms Coupled to Photonic Crystals. *Nature Photonics*, 9:326, Apr 2015. Article.

[54] A. González-Tudela, C. L. Hung, D. E. Chang, J. I. Cirac, and H. J. Kimble. Subwavelength vacuum lattices and atom–atom interactions in two-dimensional photonic crystals. *Nat. Photonics*, 9(5):320–325, 05 2015.

[55] M. Saffman, T. G. Walker, and K. Mølmer. Quantum information with Rydberg atoms. *Rev. Mod. Phys.*, 82(3):2313–2363, August 2010.

[56] Hannes Bernien, Sylvain Schwartz, Alexander Keesling, Harry Levine, Ahmed Omran, Hannes Pichler, Soonwon Choi, Alexander S. Zibrov, Manuel Endres, Markus Greiner, Vladan Vuletic, and Mikhail D. Lukin. Probing many-body dynamics on a 51-atom quantum simulator. *Nature*, 551:579, Nov 2017.

[57] Daniel Barredo, Vincent Lienhard, Sylvain de Léséleuc, Thierry Lahaye, and Antoine Browaeys. Synthetic three-dimensional atomic structures assembled atom by atom. *Nature*, 561(7721):79–82, 2018.

[58] J. R. Maze, A. Gali, E. Togan, Y. Chu, A. Trifonov, E. Kaxiras, and M. D. Lukin. Properties of Nitrogen-Vacancy Centers in Diamond: The Group Theoretic Approach. *New Journal of Physics*, 13:025025, February 2011.

[59] Bo Yan, Steven A. Moses, Bryce Gadway, Jacob P. Covey, Kaden R. A. Hazzard, Ana Maria Rey, Deborah S. Jin, and Jun Ye. Observation of Dipolar Spin-Exchange Interactions With Lattice-Confined Polar Molecules. *Nature*, 501:521, Sep 2013.

[60] Sylvain de Léséleuc, Vincent Lienhard, Pascal Scholl, Daniel Barredo, Sebastian Weber, Nicolai Lang, Hans Peter Büchler, Thierry Lahaye, and Antoine Browaeys. Observation of a symmetry-protected topological phase of interacting bosons with rydberg atoms. *Science*, 365(6455):775–780, 2019.

[61] Minh C. Tran, Andrew Y. Guo, Christopher L. Baldwin, Adam Ehrenberg, Alexey V. Gorshkov, and Andrew Lucas. The Lieb-Robinson light cone for power-law interactions. *arXiv e-prints*, page arXiv:2103.15828, March 2021.

[62] Chi-Fang Chen and Andrew Lucas. Optimal frobenius light cone in spin chains with power-law interactions, 2021.

[63] J. R. Maze, P. L. Stanwix, J. S. Hodges, S. Hong, J. M. Taylor, P. Cappellaro, L. Jiang, M. V. Gurudev Dutt, E. Togan, A. S. Zibrov, A. Yacoby, R. L. Walsworth, and M. D. Lukin. Nanoscale magnetic sensing with an individual electronic spin in diamond. *Nature*, 455(7213):644–647, October 2008.

[64] Florian Dolde, Helmut Fedder, Marcus W. Doherty, Tobias Nöbauer, Florian Rempp, Gopalakrishnan Balasubramanian, Thomas Wolf, Friedemann Reinhard, Lloyd C. L. Hollenberg, Fedor Jelezko, and Jörg Wrachtrup. Sensing electric fields using single diamond spins. *Nature Phys.*, 7(6):459–463, June 2011.

[65] Jonathon A. Sedlacek, Arne Schwettmann, Harald Kübler, Robert Löw, Tilman Pfau, and James P. Shaffer. Microwave electrometry with Rydberg atoms in a vapour cell using bright atomic resonances. *Nature Physics*, 8(11):819–824, November 2012.

[66] Christopher G. Wade, Nikola Šibalić, Natalia R. de Melo, Jorge M. Kondo, Charles S. Adams, and Kevin J. Weatherill. Real-Time Near-Field Terahertz Imaging with Atomic Optical Fluorescence. *Nat. Photonics*, 11(1):40–43, January 2017.

[67] Lincoln D. Carr, David DeMille, Roman V. Krems, and Jun Ye. Cold and ultracold molecules: Science, technology and applications. *New J. Phys.*, 11(5):055049, May 2009.

[68] A. André, A. S. Sørensen, and M. D. Lukin. Stability of Atomic Clocks Based on Entangled Atoms. *Phys. Rev. Lett.*, 92(23):230801, June 2004.

[69] Peter Høyer and Robert Špalek. Quantum Fan-out is Powerful. *Theory Comput.*, 1(1):81–103, August 2005.

[70] Miguel Aguado and Guifré Vidal. Entanglement Renormalization and Topological Order. *Phys. Rev. Lett.*, 100(7):070404, February 2008.

[71] G. Vidal. Entanglement Renormalization. *Phys. Rev. Lett.*, 99(22):220405, November 2007.

[72] G. Vidal. Class of Quantum Many-Body States That Can Be Efficiently Simulated. *Phys. Rev. Lett.*, 101(11):110501, September 2008.

[73] V. Giovannetti, S. Montangero, and Rosario Fazio. Quantum Multiscale Entanglement Renormalization Ansatz Channels. *Phys. Rev. Lett.*, 101(18):180503, October 2008.

[74] Minh C. Tran, Adam Ehrenberg, Andrew Y. Guo, Paraj Titum, Dmitry A. Abanin, and Alexey V. Gorshkov. Locality and heating in periodically driven, power-law-interacting systems. *Phys. Rev. A*, 100(5):052103, November 2019.

[75] Francisco Machado, Dominic V. Else, Gregory D. Kahanamoku-Meyer, Chetan Nayak, and Norman Y. Yao. Long-range prethermal phases of nonequilibrium matter. *Phys. Rev. X*, 10:011043, Feb 2020.

[76] Jeongwan Haah, Matthew B. Hastings, Robin Kothari, and Guang Hao Low. Quantum algorithm for simulating real time evolution of lattice Hamiltonians. *arXiv:1801.03922 [quant-ph]*, September 2018.

[77] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu. A Theory of Trotter Error. *arXiv:1912.08854 [cond-mat, physics:physics, physics:quant-ph]*, January 2020.

[78] David J. Luitz and Yevgeny Bar Lev. Emergent locality in systems with power-law interactions. *Phys. Rev. A*, 99:010105, Jan 2019.

[79] Paul Secular, Nikita Gourianov, Michael Lubasch, Sergey Dolgov, Stephen R. Clark, and Dieter Jaksch. Parallel time-dependent variational principle algorithm for matrix product states. 2019.

[80] J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi. Quantum State Transfer and Entanglement Distribution among Distant Nodes in a Quantum Network. *Phys. Rev. Lett.*, 78(16):3221–3224, 1997.

[81] Jeffrey M. Epstein and K. Birgitta Whaley. Quantum speed limits for quantum-information-processing tasks. *Phys. Rev. A*, 95(4):042314, 2017.

[82] Scott Aaronson and Alex Arkhipov. The Computational Complexity of Linear Optics. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, STOC '11, pages 333–342, New York, NY, USA, 2011. ACM.

[83] Nishad Maskara, Abhinav Deshpande, Minh C. Tran, Adam Ehrenberg, Bill Fefferman, and Alexey V. Gorshkov. Complexity phase diagram for interacting and long-range bosonic Hamiltonians. 2019.

[84] Juan Martin Maldacena. The Large N limit of superconformal field theories and supergravity. *Int. J. Theor. Phys.*, 38:1113–1133, 1999.

[85] G. S. Agarwal, R. R. Puri, and R. P. Singh. Atomic schrödinger cat states. *Phys. Rev. A*, 56:2249–2254, Sep 1997.

[86] Robert J. Lewis-Swan, Matthew A. Norcia, Julia R. K. Cline, James K. Thompson, and Ana Maria Rey. Robust spin squeezing via photon-mediated interactions on an optical clock transition. *Phys. Rev. Lett.*, 121:070403, Aug 2018.

[87] P. He, M. A. Perlin, S. R. Muleady, R. J. Lewis-Swan, R. B. Hutson, J. Ye, and A. M. Rey. Engineering spin squeezing in a 3d optical lattice with interacting spin-orbit-coupled fermions. *Phys. Rev. Research*, 1:033075, Nov 2019.

[88] Philip Richerme, Zhe-Xuan Gong, Aaron Lee, Crystal Senko, Jacob Smith, Michael Foss-Feig, Spyridon Michalakis, Alexey V. Gorshkov, and Christopher Monroe. Non-local propagation of correlations in quantum systems with long-range interactions. *Nature*, 511(7508):198–201, July 2014.

[89] Tobias J. Osborne. Efficient approximation of the dynamics of one-dimensional quantum spin systems. *Phys. Rev. Lett.*, 97:157202, Oct 2006.

[90] Stephen H. Shenker and Douglas Stanford. Black holes and the butterfly effect. *J. High Energ. Phys.*, 2014(3):67, 2014.

[91] Juan Maldacena, Stephen H. Shenker, and Douglas Stanford. A bound on chaos. *Journal of High Energy Physics*, 2016(8):106, 2016.

[92] Martin Gärttner, Justin G. Bohnet, Arghavan Safavi-Naini, Michael L. Wall, John J. Bollinger, and Ana Maria Rey. Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet. *Nature Physics*, 13(8):781–786, May 2017.

[93] Jun Li, Ruihua Fan, Hengyan Wang, Bingtian Ye, Bei Zeng, Hui Zhai, Xinhua Peng, and Jiangfeng Du. Measuring out-of-time-order correlators on a nuclear magnetic resonance quantum simulator. *Physical Review X*, 7(3), Jul 2017.

[94] Andrew Lucas. Non-perturbative dynamics of the operator size distribution in the Sachdev-Ye-Kitaev model. 2019.

[95] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[96] Brian Neyenhuis, Jiehang Zhang, Paul W. Hess, Jacob Smith, Aaron C. Lee, Phil Richerme, Zhe-Xuan Gong, Alexey V. Gorshkov, and Christopher Monroe. Observation of prethermalization in long-range interacting spin chains. *Science Advances*, 3(8), 2017.

[97] Yifan Hong and Andrew Lucas. Fast high-fidelity multi-qubit state transfer with long-range interactions. *arXiv:2009.06587 [quant-ph]*, September 2020.

[98] M. L. Mehta. *Random Matrices*. Academic Press, 3rd edition, 2004.

[99] Tianci Zhou, Shenglong Xu, Xiao Chen, Andrew Guo, and Brian Swingle. The Operator Lévy Flight: Light Cones in Chaotic Long-Range Interacting Systems. 2019.

[100] B. Nachtergaele and R. Sims. Much Ado About Something: Why Lieb-Robinson Bounds Are Useful.

[101] Marc Cheneau, Peter Barmettler, Dario Poletti, Manuel Endres, Peter Schauß, Takeshi Fukuhara, Christian Gross, Immanuel Bloch, Corinna Kollath, and Stefan Kuhr. Light-Cone-Like Spreading of Correlations in a Quantum Many-Body System. *Nature*, 481:484, Jan 2012.

[102] Nima Lashkari, Douglas Stanford, Matthew Hastings, Tobias Osborne, and Patrick Hayden. Towards the Fast Scrambling Conjecture. *Journal of High Energy Physics*, 2013(4):22, Apr 2013.

[103] M. Kliesch, C. Gogolin, and J. Eisert. *Lieb-Robinson Bounds and the Simulation of Time-Evolution of Local Observables in Lattice Systems*, page 301. 2014.

[104] Eman Hamza, Robert Sims, and Günter Stolz. Dynamical Localization in Disordered Quantum Spin Systems. *Communications in Mathematical Physics*, 315(1):215–239, Oct 2012.

[105] Peter Barmettler, Dario Poletti, Marc Cheneau, and Corinna Kollath. Propagation Front of Correlations in an Interacting Bose Gas. *Phys. Rev. A*, 85:053625, May 2012.

[106] M. B. Hastings. Light-Cone Matrix Product. *Journal of Mathematical Physics*, 50(9):095207, 2009.

[107] Ulrich Schollwöck. The Density-Matrix Renormalization Group in the Age of Matrix Product States. *Annals of Physics*, 326(1):96–192, 2011.

[108] Tilman Enss and Jesko Sirker. Light Cone Renormalization and Quantum Quenches in One-Dimensional Hubbard Models. *New J. Phys.*, 14(2):023008, 2012.

[109] M. P. Woods, M. Cramer, and M. B. Plenio. Simulating Bosonic Baths With Error Bars. *Phys. Rev. Lett.*, 115:130401, Sep 2015.

[110] M. P. Woods and M. B. Plenio. Dynamical Error Bounds for Continuum Discretisation via Gauss Quadrature Rules-a Lieb-Robinson Bound Approach. *Journal of Mathematical Physics*, 57(2):022105, 2016.

[111] M. B. Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 8:08024, August 2007.

[112] Eisert, J. and Cramer, M. and Plenio, M. B. Colloquium: Area laws for the entanglement entropy. *Rev. Mod. Phys.*, 82:277–306, Feb 2010.

[113] Zhe-Xuan Gong, Michael Foss-Feig, Fernando G. S. L. Brandão, and Alexey V. Gorshkov. Entanglement Area Laws for Long-Range Interacting Systems. *Phys. Rev. Lett.*, 119:050501, Jul 2017.

[114] Daniel Otten, Sebastian Rubbert, Jascha Ulrich, and Fabian Hassler. Universal Power-Law Decay of Electron-Electron Interactions Due to Nonlinear Screening in a Josephson Junction Array. *Phys. Rev. B*, 94:115403, Sep 2016.

[115] Minh Cong Tran, James R. Garrison, Zhe-Xuan Gong, and Alexey V. Gorshkov. Lieb-Robinson bounds on n-partite connected correlation functions. *Phys. Rev. A*, 96:052334, Nov 2017.

[116] A. I. Larkin and Y. N. Ovchinnikov. Quasiclassical Method in the Theory of Superconductivity. *Sov. Phys. JETP*, 28:1200, 1969.

[117] Alexei Kitaev and S. Josephine Suh. The soft mode in the Sachdev-Ye-Kitaev model and its gravity dual. *Journal of High Energy Physics*, 2018(5):183, May 2018.

[118] S. Michalakis. Stability of the Area Law for the Entropy of Entanglement. June 2012.

[119] J. Eisert and D. Gross. Supersonic quantum communication. *Phys. Rev. Lett.*, 102(24):240501, Jun 2009.

[120] J. Jünemann, A. Cadarso, D. Pérez-García, A. Bermudez, and J. J. García-Ripoll. Lieb-Robinson Bounds for Spin-Boson Lattice Models and Trapped Ions. *Phys. Rev. Lett.*, 111:230404, Dec 2013.

[121] Ryan Babbush, Jarrod McClean, Dave Wecker, Alán Aspuru-Guzik, and Nathan Wiebe. Chemical Basis of Trotter-Suzuki Errors in Quantum Chemistry Simulation. *Phys. Rev. A*, 91:022311, Feb 2015.

[122] *To be published*.

[123] Maksym Serbyn, Z. Papić, and D. A. Abanin. Quantum quenches in the many-body localized phase. *Phys. Rev. B*, 90:174302, Nov 2014.

[124] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009.

[125] D. L. Zhou, B. Zeng, Z. Xu, and L. You. Multiparty correlation measure based on the cumulant. *Phys. Rev. A*, 74:052110, Nov 2006.

[126] H. D. Ursell. The evaluation of gibbs' phase-integral for imperfect gases. *Math. Proc. Cambridge Philos. Soc.*, 23(6):685–697, 1927.

[127] Garrett S. Sylvester. Representations and inequalities for ising model ursell functions. *Commun. Math. Phys.*, 42(3):209–220, 1975.

[128] J. M. Hauser, W. Cassing, A. Peter, and M. H. Thoma. Connected Green function approach to ground state symmetry breaking in phi**4 in (1+1)-dimensions theory. *Z. Phys.*, A353:301–310, 1996.

[129] Peter Kopietz, Lorenz Bartosch, and Florian Schütz. *Mean-Field Theory and the Gaussian Approximation*, pages 23–52. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[130] Michael Walter, David Gross, and Jens Eisert. Multi-partite entanglement. *arXiv:1612.02437*, 2016.

[131] Silvia Pappalardi, Angelo Russomanno, Alessandro Silva, and Rosario Fazio. Multipartite entanglement after a quantum quench. *arXiv:1701.05883*, 2017.

[132] M. Hein, J. Eisert, and H. J. Briegel. Multiparty entanglement in graph states. *Phys. Rev. A*, 69:062311, Jun 2004.

[133] Robert Raussendorf, Daniel E. Browne, and Hans J. Briegel. Measurement-based quantum computation on cluster states. *Phys. Rev. A*, 68:022312, Aug 2003.

[134] Michael A. Nielsen. Cluster-state quantum computation. *Rep. Math. Phys.*, 57(1):147 – 161, 2006.

[135] Michael Kastner. Entanglement-enhanced spreading of correlations. *New J. Phys.*, 17(12):123024, 2015.

[136] Thomas Schweigler, Valentin Kasper, Sebastian Erne, Igor Mazets, Bernhard Rauer, Federica Cataldini, Tim Langen, Thomas Gasenzer, Jürgen Berges, and Jörg Schmiedmayer. Experimental characterization of a quantum many-body system via higher-order correlations. *Nature*, 545(7654):323–326, 05 2017.

[137] Tomotaka Kuwahara, Takashi Mori, and Keiji Saito. Floquet-Magnus Theory and Generic Transient Dynamics in Periodically Driven Many-Body Quantum Systems. *Ann. Phys.*, 367:96–124, April 2016.

[138] Gopikrishnan Muraleedharan, Akimasa Miyake, and Ivan H. Deutsch. Quantum computational supremacy in the sampling of bosonic random walkers on a one-dimensional lattice. *New J. Phys.*, 21:055003, 2018.

[139] M. B. Hastings. Locality in Quantum Systems. 2010.

[140] Z.-X. Gong, M. F. Maghrebi, A. Hu, M. L. Wall, M. Foss-Feig, and A. V. Gorshkov. Topological phases with long-range interactions. *Phys. Rev. B*, 93(4):041102, 2016.

[141] Dominic V. Else, Francisco Machado, Chetan Nayak, and Norman Y. Yao. An improved Lieb-Robinson bound for many-body Hamiltonians with power-law interactions. *arXiv e-prints*, page arXiv:1809.06369, Sep 2018.

[142] Jérôme Cayssol, Balázs Dóra, Ferenc Simon, and Roderich Moessner. Floquet topological insulators. *Physica Status Solidi Rapid Research Letters*, 7(1-2):101–108, Feb 2013.

[143] R. Moessner and S. L. Sondhi. Equilibration and order in quantum Floquet matter. *Nature Physics*, 13(5):424–428, Apr 2017.

[144] Dominic V. Else, Christopher Monroe, Chetan Nayak, and Norman Y. Yao. Discrete Time Crystals. *arXiv e-prints*, page arXiv:1905.13232, May 2019.

[145] Fenner Harper, Rahul Roy, Mark S. Rudner, and S. L. Sondhi. Topology and Broken Symmetry in Floquet Systems. *arXiv e-prints*, page arXiv:1905.01317, May 2019.

[146] Dmitry A. Abanin, Wojciech De Roeck, and François Huveneers. Exponentially Slow Heating in Periodically Driven Many-Body Systems. *PRL*, 115:256803, Dec 2015.

[147] Dmitry A. Abanin, Wojciech De Roeck, Wen Wei Ho, and François Huveneers. Effective Hamiltonians, prethermalization, and slow energy absorption in periodically driven many-body systems. *Physical Review B*, 95:014112, Jan 2017.

[148] Dmitry Abanin, Wojciech De Roeck, Wen Wei Ho, and François Huveneers. A Rigorous Theory of Many-Body Prethermalization for Periodically Driven and Closed Quantum Systems. *Communications in Mathematical Physics*, 354:809–827, Sep 2017.

[149] Wen Wei Ho, Ivan Protopopov, and Dmitry A. Abanin. Bounds on energy absorption and prethermalization in quantum systems with long-range interactions. *Phys. Rev. Lett.*, 120:200601, May 2018.

[150] Francisco Machado, Gregory D. Meyer, Dominic V. Else, Chetan Nayak, and Norman Y. Yao. Exponentially Slow Heating in Short and Long-range Interacting Floquet Systems. *arXiv e-prints*, page arXiv:1708.01620, Aug 2017.

[151] Takuro Matsuta, Tohru Koma, and Shu Nakamura. Improving the Lieb-Robinson Bound for Long-Range Interactions. *Ann. Henri Poincare*, 18(2):519–528, Feb 2017.

[152] S. Blanes, F. Casas, J. A. Oteo, and J. Ros. The Magnus expansion and some of its applications. *Physics Reports*, 470(5):151–238, January 2009.

[153] Marin Bukov, Luca D'Alessio, and Anatoli Polkovnikov. Universal high-frequency behavior of periodically driven systems: from dynamical stabilization to floquet engineering. *Advances in Physics*, 64(2):139–226, 2015.

[154] André Eckardt and Egidijus Anisimovas. High-frequency approximation for periodically driven quantum systems from a floquet-space perspective. *New Journal of Physics*, 17(9):093039, sep 2015.

[155] Markus Heyl, Philipp Hauke, and Peter Zoller. Quantum localization bounds Trotter errors in digital quantum simulation. *Sci. Adv.*, 5(4):eaau8342, Apr 2019.

[156] D. W. Berry, A. M. Childs, and R. Kothari. Hamiltonian simulation with nearly optimal dependence on all parameters. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 792–809, Oct 2015.

[157] Arijeet Pal and David A. Huse. Many-body localization phase transition. *Phys. Rev. B*, 82:174411, Nov 2010.

[158] Ryan Babbush, Nathan Wiebe, Jarrod McClean, James McClain, Hartmut Neven, and Garnet Kin-Lic Chan. Low-Depth Quantum Simulation of Materials. *Phys. Rev. X*, 8(1):011044, March 2018.

[159] Anders W. Sandvik. Computational Studies of Quantum Spin Systems. In Adolfo Avella and Ferdinando Mancini, editors, *American Institute of Physics Conference Series*, volume 1297 of *American Institute of Physics Conference Series*, pages 135–338, Nov 2010.

[160] Masuo Suzuki. General theory of fractal path integrals with applications to many-body theories and statistical physics. *Journal of Mathematical Physics*, 32(2):400–407, February 1991.

[161] Andrew M. Childs and Nathan Wiebe. Hamiltonian Simulation Using Linear Combinations of Unitary Operations. *QIC*, 12(11-12).

[162] Earl Campbell. A random compiler for fast Hamiltonian simulation. *Phys. Rev. Lett.*, 123(7):070503, August 2019.

[163] Andrew M. Childs, Aaron Ostrander, and Yuan Su. Faster quantum simulation by randomization. *Quantum*, 3:182, September 2019.

[164] Andreas V. Kuhlmann, Julien Houel, Arne Ludwig, Lukas Greuter, Dirk Reuter, Andreas D. Wieck, Martino Poggio, and Richard J. Warburton. Charge noise and spin noise in a semiconductor quantum device. *Nature Physics*, 9(9):570–575, September 2013.

[165] Paolo Zanardi. Symmetrizing evolutions. *Physics Letters A*, 258(2-3):77–82, 1999.

[166] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical Decoupling of Open Quantum Systems. *Phys. Rev. Lett.*, 82(12):2417–2421, March 1999.

[167] P. Facchi, D. A. Lidar, and S. Pascazio. Unification of dynamical decoupling and the quantum Zeno effect. *Phys. Rev. A*, 69(3):032314, March 2004.

[168] Kaveh Khodjasteh and Daniel A. Lidar. Rigorous bounds on the performance of a hybrid dynamical-decoupling quantum-computing scheme. *Phys. Rev. A*, 78:012355, Jul 2008.

[169] Lorenza Viola, Seth Lloyd, and Emanuel Knill. Universal Control of Decoupled Quantum Systems. *Phys. Rev. Lett.*, 83(23):4888–4891, December 1999.

[170] Hui Khoon Ng, Daniel A. Lidar, and John Preskill. Combining dynamical decoupling with fault-tolerant quantum computation. *Phys. Rev. A*, 84(1):012305, July 2011.

[171] Daniel Burgarth, Paolo Facchi, Giovanni Gramegna, and Saverio Pascazio. Generalized Product Formulas and Quantum Control. *J. Phys. A: Math. Theor.*, 52(43):435301, October 2019.

[172] K. Stannigel, P. Hauke, D. Marcos, M. Hafezi, S. Diehl, M. Dalmonte, and P. Zoller. Constrained Dynamics via the Zeno Effect in Quantum Simulation: Implementing Non-Abelian Lattice Gauge Theories with Cold Atoms. *Phys. Rev. Lett.*, 112(12):120406, March 2014.

[173] Rahul Nandkishore and David A. Huse. Many-Body Localization and Thermalization in Quantum Statistical Mechanics. *Annual Review of Condensed Matter Physics*, 6(1):15–38, 2015.

[174] David J. Luitz, Nicolas Laflorencie, and Fabien Alet. Many-body localization edge in the random-field Heisenberg chain. *Phys. Rev. B*, 91(8):081103, February 2015.

[175] Stephen P. Jordan, Keith S. M. Lee, and John Preskill. Quantum Computation of Scattering in Scalar Quantum Field Theories. *arXiv:1112.4833 [hep-th, physics:quant-ph]*, January 2019.

[176] Stephen P. Jordan, Keith S. M. Lee, and John Preskill. Quantum Algorithms for Quantum Field Theories. *Science*, 336(6085):1130–1133, June 2012.

[177] Philipp Hauke, David Marcos, Marcello Dalmonte, and Peter Zoller. Quantum simulation of a lattice Schwinger model in a chain of trapped ions. *Phys. Rev. X*, 3(4):041018, November 2013.

[178] Erez Zohar, J. Ignacio Cirac, and Benni Reznik. Cold-Atom Quantum Simulator for SU(2) Yang-Mills Lattice Gauge Theory. *Phys. Rev. Lett.*, 110(12):125304, March 2013.

[179] Zohreh Davoudi, Mohammad Hafezi, Christopher Monroe, Guido Pagano, Alireza Seif, and Andrew Shaw. Towards analog quantum simulations of lattice gauge theories with trapped ions. *Phys. Rev. Research*, 2(2):023015, April 2020.

[180] E. A. Martinez, C. A. Muschik, P. Schindler, D. Nigg, A. Erhard, M. Heyl, P. Hauke, M. Dalmonte, T. Monz, P. Zoller, and R. Blatt. Real-time dynamics of lattice gauge theories with a few-qubit quantum computer. *Nature*, 534(7608):516–519, June 2016.

[181] N. Klco, E. F. Dumitrescu, A. J. McCaskey, T. D. Morris, R. C. Pooser, M. Sanz, E. Solano, P. Lougovski, and M. J. Savage. Quantum-Classical Computation of Schwinger Model Dynamics using Quantum Computers. *Phys. Rev. A*, 98(3):032331, September 2018.

[182] Natalie Klco, Jesse R. Stryker, and Martin J. Savage. SU(2) non-Abelian gauge field theory in one dimension on digital quantum computers. *Phys. Rev. D*, 101(7):074512, April 2020.

[183] Bipasha Chakraborty, Masazumi Honda, Taku Izubuchi, Yuta Kikuchi, and Akio Tomiya. Digital Quantum Simulation of the Schwinger Model with Topological Term via Adiabatic State Preparation. *arXiv:2001.00485 [cond-mat, physics:hep-lat, physics:hep-ph, physics:hep-th, physics:quant-ph]*, February 2020.

[184] Alexander F. Shaw, Pavel Lougovski, Jesse R. Stryker, and Nathan Wiebe. Quantum Algorithms for Simulating the Lattice Schwinger Model. *arXiv:2002.11146 [hep-lat, physics:nucl-th, physics:quant-ph]*, February 2020.

[185] John Kogut and Leonard Susskind. Hamiltonian formulation of Wilson's lattice gauge theories. *Phys. Rev. D*, 11(2):395–408, January 1975.

[186] Sidney Coleman. More about the massive schwinger model. *Annals of Physics*, 101(1):239 – 267, 1976.

[187] C. J. Hamer, Zheng Weihong, and J. Oitmaa. Series expansions for the massive Schwinger model in Hamiltonian lattice theory. *Phys. Rev. D*, 56(1):55–67, July 1997.

[188] T. M. R. Byrnes, P. Sriganesh, R. J. Bursill, and C. J. Hamer. Density matrix renormalization group approach to the massive Schwinger model. *Phys. Rev. D*, 66(1):013002, July 2002.

[189] Boye Buyens, Jutho Haegeman, Karel Van Acoleyen, Henri Verschelde, and Frank Verstraete. Matrix product states for gauge field theories. *Phys. Rev. Lett.*, 113:091601, Aug 2014.

[190] Stefan Kühn, J. Ignacio Cirac, and Mari-Carmen Bañuls. Quantum simulation of the Schwinger model: A study of feasibility. *Phys. Rev. A*, 90(4):042305, October 2014.

[191] Elisa Ercolessi, Paolo Facchi, Giuseppe Magnifico, Saverio Pascazio, and Francesco V. Pepe. Phase Transitions in Z_{n} Gauge Models: Towards Quantum Simulations of the Schwinger-Weyl QED. *Phys. Rev. D*, 98(7):074503, October 2018.

[192] Siu A. Chin. Multi-product splitting and Runge-Kutta-Nyström integrators. *Celest Mech Dyn Astr*, 106(4):391–406, April 2010.

[193] Jens Eisert, Mauritz van den Worm, Salvatore R. Manmana, and Michael Kastner. Breakdown of Quasilocality in Long-Range Quantum Lattice Models. *Phys. Rev. Lett.*, 111(26):260401, December 2013.

[194] P. Hauke and L. Tagliacozzo. Spread of Correlations in Long-Range Interacting Quantum Systems. *Phys. Rev. Lett.*, 111(20):207202, November 2013.

[195] Andrew Y. Guo, Minh C. Tran, Andrew M. Childs, Alexey V. Gorshkov, and Zhe-Xuan Gong. Signaling and Scrambling with Strongly Long-Range Interactions. *Phys. Rev. A*, 102(1):010401, July 2020.

[196] David Poulin, Matthew B. Hastings, Dave Wecker, Nathan Wiebe, Andrew C. Doberty, and Matthias Troyer. The trotter step size required for accurate quantum simulation of quantum chemistry. *Quantum Info. Comput.*, 15(5-6):361–384, April 2015.

[197] Mario Motta, Erika Ye, Jarrod R. McClean, Zhendong Li, Austin J. Minnich, Ryan Babbush, and Garnet Kin-Lic Chan. Low rank representations for quantum simulation of electronic structure. *arXiv:1808.02625 [physics, physics:quant-ph]*, August 2018.

[198] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan. Quantum computational chemistry. *Rev. Mod. Phys.*, 92:015003, Mar 2020.

[199] Cristina Cirstoiu, Zoe Holmes, Joseph Iosue, Lukasz Cincio, Patrick J Coles, and Andrew Sornborger. Variational fast forwarding for quantum simulation beyond the coherence time. *arXiv preprint arXiv:1910.04292*, 2019.

[200] Burak Şahinoğlu and Rolando D. Somma. Hamiltonian simulation in the low energy subspace. *arXiv:2006.02660 [quant-ph]*, June 2020.

[201] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*, page 36. Cambridge University Press, New York, NY, 2004.

[202] Daniel A. Klain and Gian-Carlo Rota. *Introduction to Geometric Probability*, pages 46–47. Cambridge University Press, New York, NY, 1997.

[203] Quaintance J. and H. W. Gould. *Combinatorial Identities for Stirling Numbers The Unpublished Notes of H W Gould*. World Scientific Publishing Co., Singapore.

[204] Jesse R. Stryker. Oracles for Gauss's law on digital quantum computers. *Phys. Rev. A*, 99(4):042301, April 2019.

[205] Per Christian Moan and Jitse Niesen. Convergence of the magnus series. *Foundations of Computational Mathematics*, 8(3):291–301, Jun 2008.

[206] Ana Arnal, Fernando Casas, and Cristina Chiralt. A general formula for the Magnus expansion in terms of iterated integrals of right-nested commutators. *arXiv:1710.10851 [math-ph]*, October 2017.